Characterisation and Control of 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate Synthase from *Geobacillus sp*

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Science in Biochemistry

at the

University of Canterbury

by

Mohamad Othman

October 2014

Abstract

3-Deoxy-D-*arabino* heptulosonate 7-phosphate synthase (DAH7PS) catalyses the first step of the shikimate pathway, responsible for the biosynthesis of aromatic amino acids. This pathway is found in microorganisms, plants and apicomplexan parasites and its absence in mammals makes it a viable target for antimicrobial drug design. DAH7PS enzymes differ in the regulatory machinery that decorates the catalytic $(\beta/\alpha)_8$ barrel. Some DAH7PS enzymes are fused to chorismate mutase (CM), another enzyme in the shikimate pathway. This fusion protein is allosterically regulated by chorismate (CA) or prephenate (PA), the precursor of tyrosine and phenylalanine. It has been suggested that DAH7PS enzymes evolved these extensions to the core barrel for the sole purpose of regulation.

Geobacillus sp DAH7PS (*Gsp*DAH7PS^{WT}) is a thermophilic type Iβ DAH7PS enzyme with an N-terminal CM domain fused through a linker region. This thesis describes the functional characterisation work carried out on *Gsp*DAH7PS^{WT}, in attempt to help determine how DAH7PS enzymes evolved such diverse methods of regulation.

Chapter 2 describes the functional characterisation work carried out on the catalytic and regulatory domains of *Gsp*DAH7PS^{WT}. The enzyme demonstrated both DAH7PS and CM activities with the DAH7PS domain determined to be metal dependent and most activated by Cd^{2+}. PA completely inhibited the catalytic activity of *Gsp*DAH7PS^{WT}, and AUC demonstrated an equilibrium exists between the dimeric and tetrameric quaternary states of the enzyme in solution.

Chapter 3 describes the domain truncation of *Gsp*DAH7PS^{WT} carried out at the linker region in order to obtain two separate protein domains, the catalytic domain lacking the N-terminal domain (*Gsp*DAH7PS^{DAH7PS}) and the regulatory domain without the catalytic domain (*Gsp*DAH7PS^{CM}). Both variants were fully characterised, and information obtained from each domain was compared to the respective catalytic and regulatory domains of the wild-type enzyme, which was also characterised. Like *Gsp*DAH7PS^{WT}, *Gsp*DAH7PS^{DAH7PS}

showed greatest activation in the presence of Cd^{2+}, with other metals having varying effects on activation rates and stability of the enzyme. Both truncated variants followed Michaelis-Menten kinetics where *Gsp*DAH7PS^{DAH7PS} was found to be more active than *Gsp*DAH7PS^{WT} and unaffected by PA, whereas *Gsp*DAH7PS^{CM} was a less efficient catalyst than the CM domain of *Gsp*DAH7PS^{WT}. AUC demonstrated that in solution an equilibrium occurs between the monomeric and tetrameric oligomeric states of *Gsp*DAH7PS^{DAH7PS}.

Chapter 4 summarises the findings of the thesis along with future directions of this research, combining the results obtained and expanding upon them. It is concluded that the catalytic regulatory CM domain supports both protein structure and allosteric regulation of *Gsp*DAH7PS^{WT}

Acknowledgements

I really cannot believe that I have finally got to this point. This definitely has been the most challenging part of my life so far, but in return I have learnt so much more about myself and my field of study, with a learning experience unlike anything I have encountered before.

First of all I would like to thank my wonderful supervisor Emily Parker. Thank you for all the help, without you this thesis would not have been possible. Thank you for always being there for me throughout the good and bad times. You will always be someone that I look up to and always wonder how you manage to be so amazing at everything you do.

Thank you to Ali Reza Nazmi who co-supervised me in the lab and taught me the necessary basics to getting started.

Thank you to the Parker group, you guys are a great bunch. I would like to give a special thanks to Penel Cross, Michael Hunter, Tyler Clarke, Ryu Toyama, Michael Weusten, Logan Heyes, Nicky Blackmore, Gerd Mittlestaedt, Vicky Zhang, Dmitri Joseph, Emma Livingstone, Thomas Cotton, Tammie Cookson, Effie Fan, Sarah Wilson-Coutts, Gert-Jan Moggre, Sebastian Reichau, and Eric Lang who all played a role in helping me get through this thesis.

Thank you to my amazing parents, for all the love, support and all the sacrifices you have made to ensure I got the opportunity to get to this point. To my siblings, just like all other siblings we have had our ups and downs, and in the end we have always stuck together and have been there for one another as we will always continue to do so. Thank you for all of your support, you are my everything and I love you all.

Contents

Abbreviations

A5P arabinose 5-phosphate

Amp ampicillin

***Ape*DAH7PS** *Aeropyrum pernix* DAH7PS

Arg arginine

Asp aspartic acid

AUC analytic ultracentrifugation

***Bsu*DAH7PS** *Bacillus subtilis* DAH7PS

BTP 1,3-(tris(hydroxymethyl)-methylamino)propane

CA chorismate

Cam chloramphenicol

CD circular dichroism

CM chorismate mutase

Cys cysteine

DAH7P 3-deoxy-D-*arabino*-heptulosonate 7-phosphate

DAH7PS 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase

DNA deoxyribonucleic acid

dNTP deoxyribonucleotide triphosphate

DPA dipicolinic acid

DSF differential scanning fluorimetry

DTT dithiothreitol

E4P erythrose 4-phosphate

***Eco*DAH7PS** *Escherichia coli* DAH7PS

EDTA ethylenediaminetetraacetic acid

FL ferredoxin-like

G3P glycerol 3-phosphate

***Gsp*DAH7PS^{WT}** *Geobacillus sp* DAH7PS

***Gsp*DAH7PS^{DAH7PS}** *Geobacillus sp* DAH7PS truncated variant

***Gsp*DAH7PS^{CM}** *Geobacillus sp* CM truncated variant

Glu glutamic acid

GST glutathione S-transferase

His histidine

pI isoelectric point

IPTG isopropyl-1-thiol-β-D-galactopyranoside

Kan kanamycin

KDO 3-deoxy-D-*manno*-2-octulosonic acid

KDO8P 3-deoxy-D-*manno*-octulosonate 8-phosphate

KDO8PS 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase

LB lysogeny broth

***Lmo*DAH7PS** *Listeria monocytogenes* DAH7PS

Lys lysine

MBP maltose binding protein

***Mtu*DAH7PS** *Mycobacterium tuberculosis* DAH7PS

***Mtu*CM** *Mycobacterium tuberculosis* CM

MWCO molecular weight cut-off

ORF open reading frame

PA prephenate

PAGE polyacrylamide gel electrophoresis

PCR polymerase chain reaction

PDB Protein Data Bank

PEP phosphoenolpyruvate

***Pfu*DAH7PS** *Pyrococcus furiosus* DAH7PS

***Pgi*DAH7PS** *Porphyromonas gingivalis* DAH7PS

Phe phenylalanine

pI isoelectric point

Pro proline

RMSD root-mean-square deviation

SAXS small angle X-ray scattering

***Sce*DAH7PS** *Saccharomyces cerevisiae* DAH7PS

SDS sodium dodecyl sulfate

SEC size-exclusion chromatography

Ser serine

SOC super optimal broth

Spec spectinomycin

TAE tris base-acetic acid-EDTA

TE tris-EDTA

TEV Tobacco Etch Virus

Thr threonine

TIM triosephosphate isomerase

***Tma*DAH7PS** *Thermotoga maritima* DAH7PS

Trp tryptophan

Tyr tyrosine

UV ultra-violet

WT wild-type

List of figures

List of tables

Chapter 1

Introduction

1.1 The shikimate pathway

Different organisms differ in their ability to synthesise amino acids that are essential for them to thrive. Amino acids are of great importance, serving as building blocks of proteins and as intermediates in metabolic pathways.[1] Animals, unlike microorganisms, are unable to produce all the necessary amino acids, and therefore these amino acids need to be acquired by means of diet. Humans are only able to produce half of the required amino acids to support metabolism and therefore depend on food to supply the other half.[2] Tyrosine (Tyr), phenylalanine (Phe) and tryptophan (Trp) are three amino acids not synthesised in animals that are synthesised via the shikimate pathway. This pathway operates exclusively in microorganisms, plants and apicomplexan parasites.[3]

The shikimate pathway consists of a series of seven enzyme-catalysed reactions leading to the production of chorismate (CA), at which point branching of the shikimate pathway occurs towards a variety of aromatic end products (Figure 1.1).

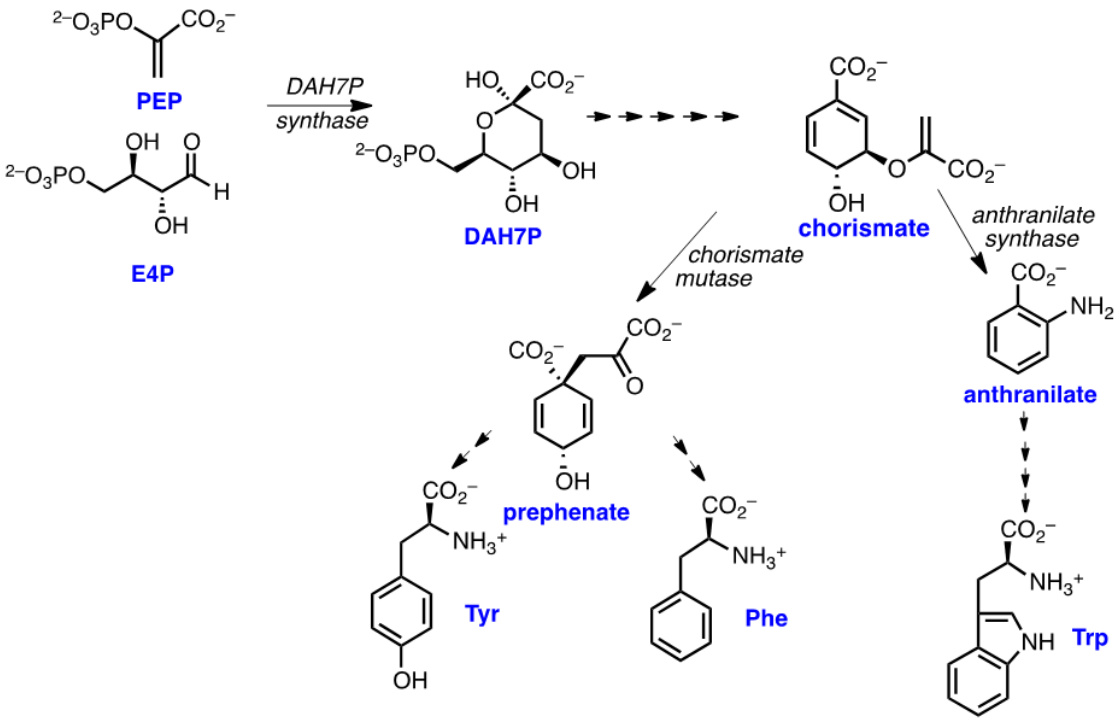


Figure 1.1: Overview of shikimate pathway focusing on the key reactions involved in this project, including synthesis of chorismate prior to branching of the pathway and formation of the vital amino acids Tyr, Phe and Trp. Dashed arrows represent reactions that involve multiple steps.

The shikimate pathway was elucidated by Davis and Sprinson along with their associates in the 1950s.[4] The absence of the shikimate pathway in mammals has made the enzymes involved possible targets for herbicides and antimicrobial drug design.[5-7] This means that inhibitors could be synthesised to target and halt the shikimate pathway in pathogenic organisms while having no effect on the animals that are being treated.

1.2 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase

The first step of the shikimate pathway is an ordered sequential aldol condensation reaction catalysed by the enzyme 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS).[8] In this reaction, phosphoenolpyruvate (PEP) is coupled with a four carbon phosphorylated sugar, D-erythrose 4-phosphate (E4P), to form 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAH7P) and inorganic phosphate (Figure 1.1).[9]

As a way of controlling the rate of production of shikimate pathway products, DAH7PS is often regulated by a feedback inhibition mechanism by the end products or intermediates

of the pathway. Different organisms have been shown to employ different strategies for regulating the biosynthetic activity of the shikimate pathway. DAH7PS enzymes from some sources are inhibited by Tyr, Phe or Trp, whereas others are inhibited by combinations of the aromatic amino acids.[10,11] In some bacteria, DAH7PS is found to be fused to chorismate mutase (CM), a chorismate utilising enzyme on the branch to Phe and Tyr.[12] The DAH7PS domain of this fusion protein is allosterically regulated by chorismate or prephenate (PA), the precursor to Tyr and Phe.[13]

1.3 Different types of DAH7PS enzymes

All DAH7PS enzymes share the same core $(\beta/\alpha)_8$ barrel fold.[14] However, there is considerable variation in the different small domain decorations that are attached to this catalytic barrel.[15] These structural additions are responsible for the allosteric regulation of the enzyme through the binding of the biosynthetic pathway end products, with different enzymes demonstrating different mechanisms of inhibition.[16-18] For some DAH7PS enzymes, the binding of inhibitors alter the enzyme's conformation or dynamics leading to a change in the active site and thus preventing the substrates from binding.[16]

The DAH7PS enzymes have been divided into two classes (type I and type II according to Jensen) depending on their molecular weight and amino acid sequence.[19] Type I is further subdivided into Iα and Iβ subfamilies based solely on sequence similarity. There is a low sequence similarity between individual members of subtypes Iα and Iβ, (approximately 18 %) suggesting limited homology between the two subfamilies (Figure 1.2).[18] Type Iα DAH7PS enzymes are generally larger (approximately 40 kDa) than type Iβ DAH7PS enzymes (approximately 30 kDa). Both subtypes however are smaller than type II DAH7PS (>54 kDa) enzymes.[20,21]

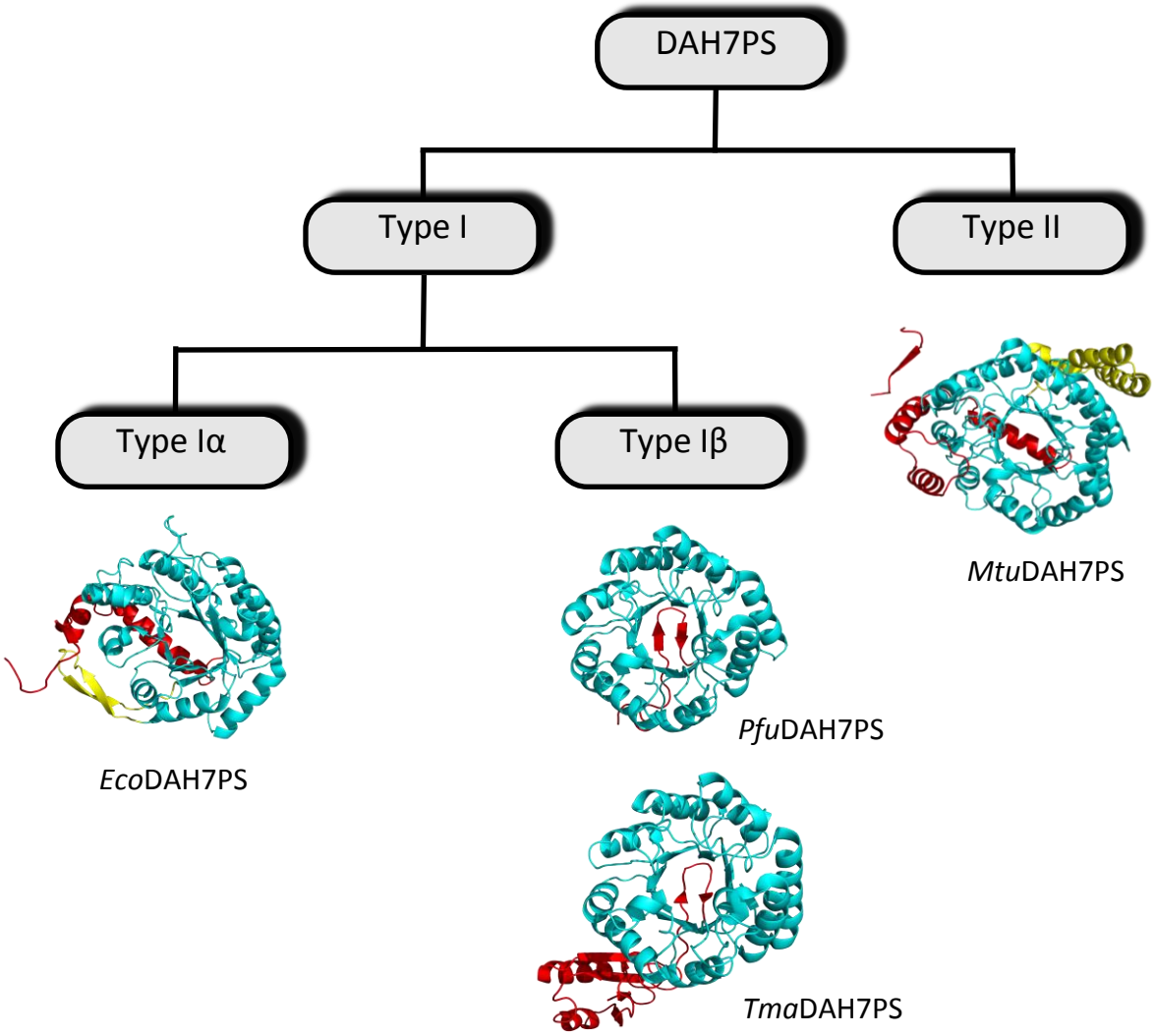


Figure 1.2: The different DAH7PS families showing the monomer fold of previously characterised DAH7PS enzymes. The core $(\beta/\alpha)_8$ barrel is shown in cyan with internal and external extensions shown in yellow and red, respectively.

While subtypes Iα and Iβ belong to the same DAH7PS family and catalyse the same reaction, based upon sequence identity, subtype Iβ DAH7PS enzymes are functionally and structurally more closely related to 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase (KDO8PS) enzymes ($I\beta_K$) than to the DAH7PS enzymes of subtype Iα.[20]

DAH7PS and KDO8PS are two functionally unrelated enzymes with low sequence similarity, yet they have features in common such as mechanism and structure.[22] KDO8PS catalyses a reaction analogous to that of DAH7PS, condensing D-arabinose 5-phosphate (A5P) instead of E4P with PEP to form 3-deoxy-D-*manno*-octulosonate 8-phosphate (KDO8P) and inorganic phosphate.[23] This is an important reaction that produces 3-deoxy-D-*manno*-octulosonate

(KDO), a product that is necessary for the biosynthesis of lipopolysaccharides that are required for the growth of Gram negative bacteria (Figure 1.3).[24]

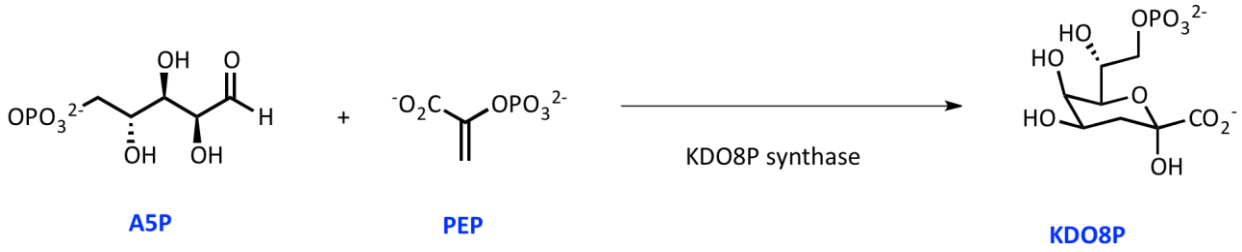


Figure 1.3: Condensation of PEP and A5P catalysed by KDO8P synthase.

1.4 Catalytic mechanism

Numerous details have been obtained about the reaction mechanism of DAH7PS through extensive studies (Figure 1.4). Originally it was thought that the reaction between PEP and E4P occurred through a ping-pong mechanism.[25] This was later disproven by another study showing that the reaction followed an ordered sequential kinetic mechanism.

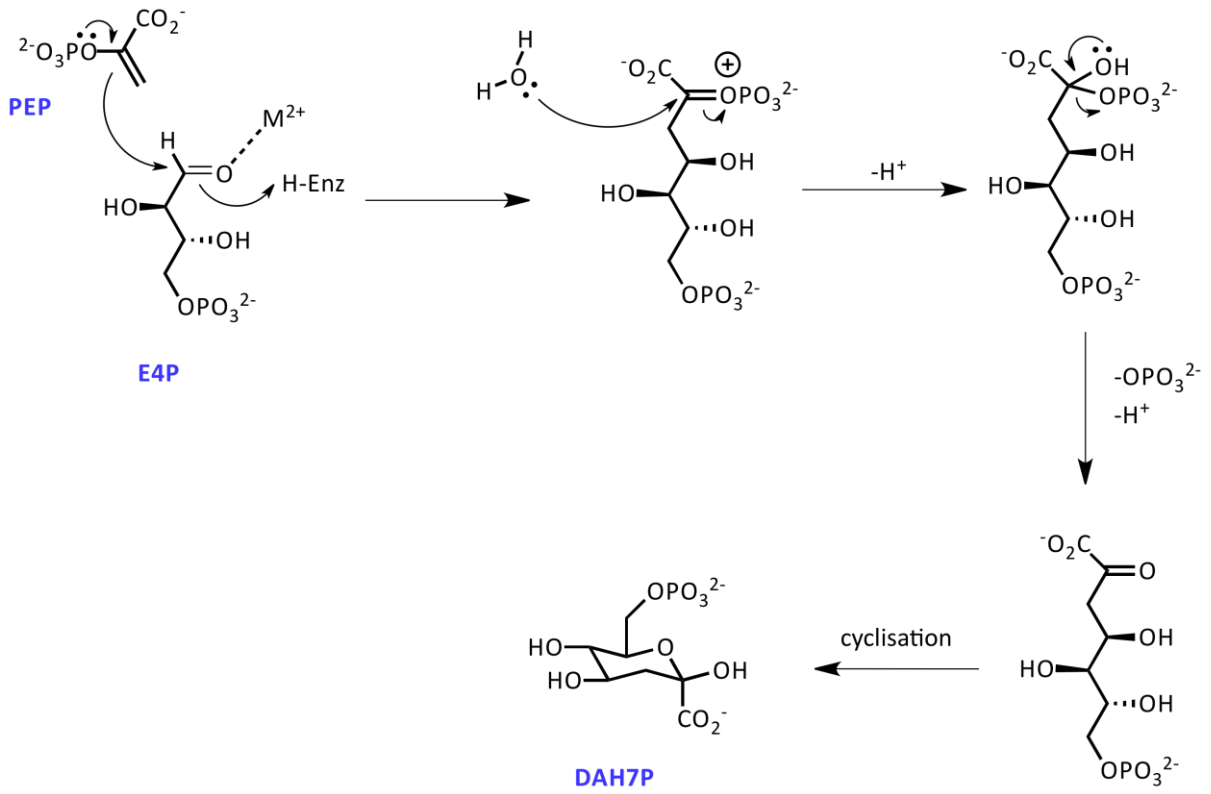


Figure 1.4: DAH7PS catalysed condensation mechanism of PEP and E4P, to form DAH7P.

This mechanism proceeds in a stereospecific manner where PEP binds to the enzyme first, followed by E4P,[26] with the *si* face of PEP coupling with the *re* face of E4P to produce DAH7P (Figure 1.5).[27,28] This is followed by water attack on the oxocarbenium ion, phosphate loss, and finally DAH7P release.[26]

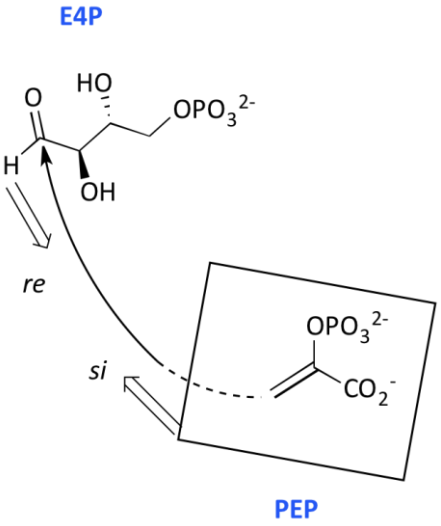
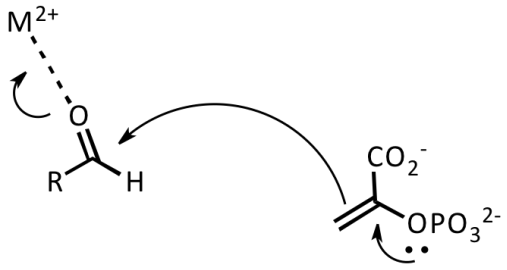


Figure 1.5: DAH7PS catalysed stereospecific reaction of PEP and E4P.

The loss of phosphate from PEP occurs through the cleavage of the C-O bond instead of the P-O bond of PEP. This is unusual due to the high energy contained in the phosphate ester bond that could be utilised by the enzyme catalysed reaction.[26] The oxygen at C2 of DAH7P is therefore obtained from water which attacks at C2 of PEP.

It is thought that the first step of the reaction, where C3 of PEP attacks C1 of E4P, is activated by the coordination of its carbonyl group to an essential divalent metal ion, to form the oxocarbenium ion intermediate (Figure 1.6).[29] An active site water is then believed to attack the *re* face of the intermediate to form a phosphate hemiketal intermediate. This intermediate loses its phosphate through C-O bond cleavage to form DAH7P. This seven carbon sugar is then thought to be released from the enzyme in an acyclic form that spontaneously equilibrates to a more stable cyclic pyranose form.[30]



Lewis Acid mechanism

Figure 1.6: Mechanism showing the proposed role of the metal ion in the active site of DAH7PS.

1.5 Metal activation

While both metal dependent and metal independent KDO8PS enzymes have been characterised, all DAH7PS enzymes characterised to date have been dependent on a divalent metal ion for catalysis. [31-35]

The presence of the metal chelator ethylenediaminetetraacetic acid (EDTA) causes most DAH7PS enzymes characterised to date to lose activity. Activity is restored upon the addition of divalent metal ions.[7,12,13,36] Sequence alignment studies show all known DAH7PS enzymes contain a set of conserved metal binding residues (Cys, His, Glu and Asp) that contribute to the coordination of the metal ion.[37,38] Metal independent KDO8PS enzymes have been shown also to have the conserved metal binding residues, however Cys has been substituted to Asn. An example of this is observed in *Escherichia coli* KDO8PS, a metal independent enzyme, which contains an Asn instead of Cys as part of the conserved metal binding residues.[22] A study was carried out where Asn was substituted with Cys, resulting in a variant with increased enzymatic activity in the presence of divalent metal.

Whereas all DAH7PS enzymes require metal for enzyme catalysis, a number of different divalent metal ions including Mn^{2+}, Cd^{2+} and Zn^{2+}, are able to activate the enzyme. Some metal ions are found to be more activating than others, depending on the source of the enzyme.[6,13,15,39,40] This variation in metal selection is also the case in the different isozymes of the same organism. An example of this involves the Tyr-regulated DAH7PS isozyme of

Saccharomyces cerevisiae (*Sce*DAH7PS) which is most activated by Cd^{2+}, whereas the Phe-regulated isozyme demonstrates the highest catalytic activity in the presence of Mn^{2+}.[31]

1.6 Structure of DAH7PS

Obtaining crystal structures of DAH7PS enzymes can provide a great amount of detail about the enzyme that is unobtainable with functional studies. The crystal structure allows for visual inspection, helping to determine the way in which the enzyme is folded, the position of monomeric units and how they interact with one another.[7,37,41]

Type 1α

The genome of *E. coli* encodes for three DAH7PS isozymes (*Eco*DAH7PS). All three isozymes are classified as 1α subfamily of DAH7PS enzymes, with each being specifically inhibited by only one aromatic amino acid from the shikimate pathway. Sequence analysis determined that the three isozymes are similar in molecular weight, sequence and polypeptide chain length.[37]

The crystal structure of the isozyme inhibited by Phe (*Eco*DAH7PS(Phe)), was the first structure of a DAH7PS enzyme to be solved.[37] The enzyme crystallised as a tetramer consisting of a dimer of two tight dimers (Figure 1.7). Association of the subunits in the tight dimer is significantly enhanced due to the formation of a three stranded, antiparallel β sheet involving the N-terminal β0 strand of one monomer and the extended β6a- β6b strands of the adjacent chain.

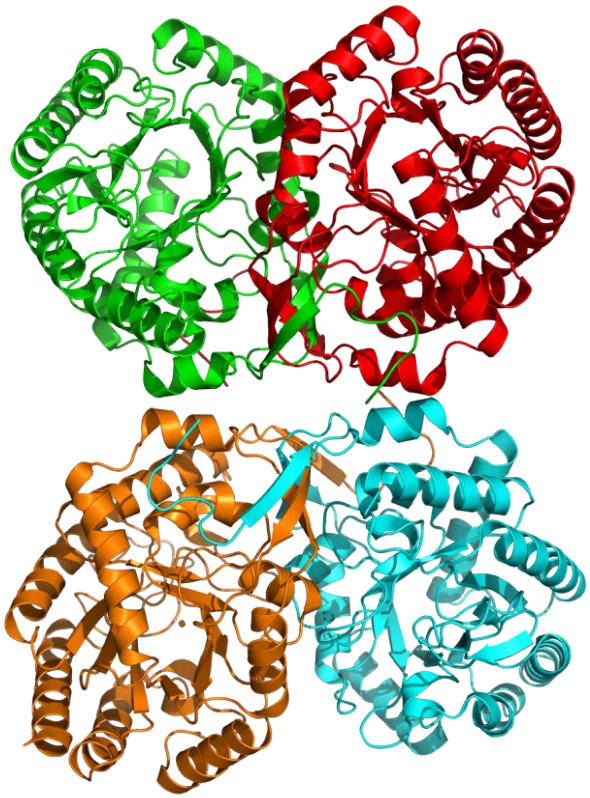


Figure 1.7: Tetrameric structure of *Eco*DAH7PS (PDB code 1KFL) with different colours representing individual monomers.

Each chain of *Eco*DAH7PS(Phe) contains a $(\beta/\alpha)_8$ triosephosphate isomerase (TIM)-barrel and this structure is observed for all other DAH7PS enzymes characterised to date.[7,11,13,31] The active site of this fold is located at the C-terminal end of the barrel.[42] Extended loops that connect helices and strands at the C-terminal (β-α loops) harbour the residues that make up the active site for the binding of the metal ion and the substrates PEP and E4P. These loops are longer than those at the N-terminal end of the barrel (α-β loops).

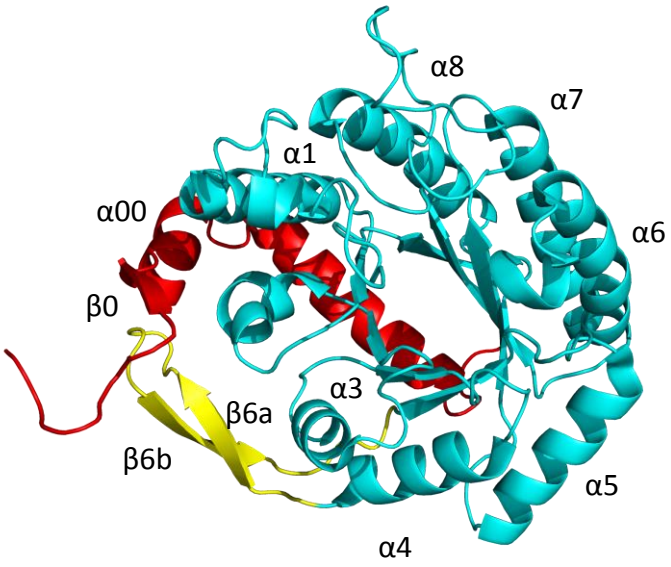


Figure 1.8: Monomeric structure of *Eco*DAH7PS (PDB code 1KFL) showing the core $(\beta/\alpha)_8$ barrel (cyan), with the N-terminal extension (red) and internal extension (yellow).

The crystal structure of *Eco*DAH7PS(Phe) also shows that each subunit is enhanced by a 53 residue long N-terminal extension, along with a two-stranded antiparallel β-sheet β6a/β6b inserted between the α5 and β6 strands of the barrel (Figure 1.8).[37] These additions to the core barrel have been shown to be involved in feedback regulation of the enzyme and provide an allosteric binding site for Phe.[8,43]

Type Iβ

The crystal structure of *Thermotoga maritima* DAH7PS (*Tma*DAH7PS) was the first structure to be solved for a Iβ DAH7PS enzyme.[41] The enzyme, like subtype Iα DAH7PS enzymes, contains the core $(\beta/\alpha)_8$ barrel with a similar active site architecture. It was shown that *Tma*DAH7PS forms a tetramer made up of two dimers associating through helices α5-α8 and loops β5-α5, β6-α6, and β7-α7.

Each core barrel has a two stranded hairpin and an N-terminal ferredoxin-like (FL) domain consisting of a distinctive βαββαβ fold connected to the barrel though a flexible linker, which was determined to be involved with feedback regulation of the enzyme (Figure 1.9).[16]

The FL domain is notably similar to the ACT domain, found fused to a number of other enzymes involved in amino acid biosynthesis, for the purpose of ligand binding and regulation.[44] Recognised as a recurring motif among a diverse group of proteins in 1999,[45] the ACT domain was named after the first letters of three proteins, aspartate kinase, chorismate mutase and tyrA (prephenate dehydrogenase), in which it was originally identified.

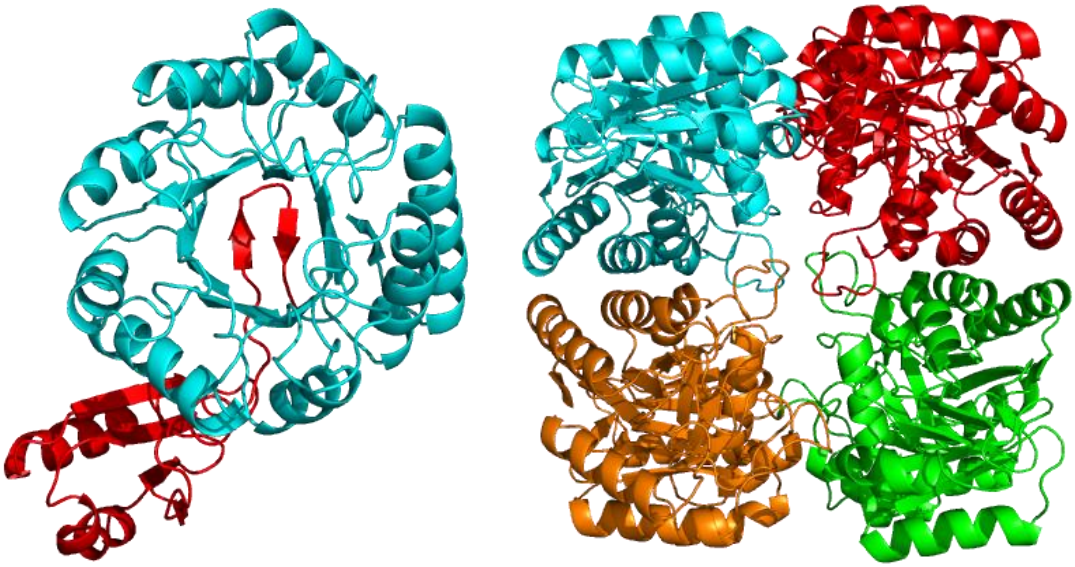


Figure 1.9: Monomeric (left) and tetrameric (right) structure of *Tma*DAH7PS (PDB code 1RZM). Monomeric structure shows the core $(\beta/\alpha)_8$ barrel (cyan) with the N-terminal extension (red). The tetrameric structure is shown with different colours representing individual monomers.

Another Iβ enzyme is that of *Pyrococcus furiosus* DAH7PS (*Pfu*DAH7PS).[14] This enzyme also crystallises as a tetramer, with each subunit consisting of the $(\beta/\alpha)_8$ barrel with the two stranded hairpin similar to the subunit of *Tma*DAH7PS. *Pfu*DAH7PS lacks the ferredoxin-like of *Tma*DAH7PS, however its crystal structure consists of the same tetrameric structure as *Tma*DAH7PS (Figure 1.10).

Similar to *Pfu*DAH7PS, *Aeropyrum pernix* DAH7PS (*Ape*DAH7PS) lacks any added external or internal decorations to the $(\beta/\alpha)_8$ barrel. The crystal structure of *Ape*DAH7PS shows that the enzyme forms the same tightly associated tetramer as *Pfu*DAH7PS and *Tma*DAH7PS (Figure 1.10).[15]

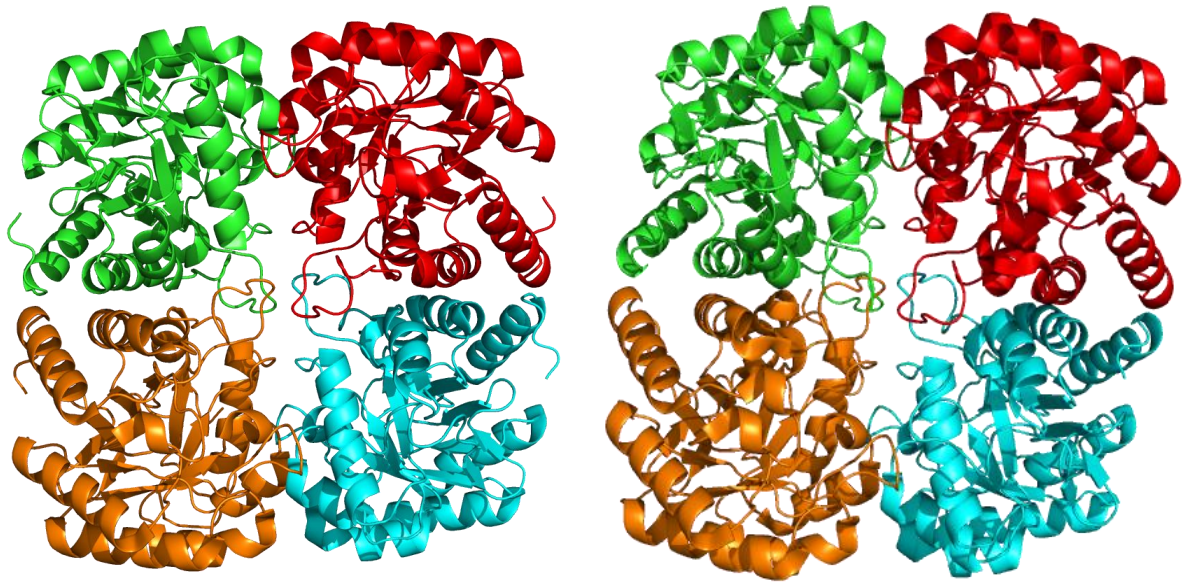


Figure 1.10: Tetrameric crystal structures of *Pfu*DAH7PS (PDB code 1ZCO) (left) and *Ape*DAH7PS (PDB code 1VS1) (right) with different colours representing individual monomers.

Type II

The crystal structure of *Mycobacterium tuberculosis* DAH7PS (*Mtu*DAH7PS) has been solved and is the, being the first only type II DAH7PS enzyme structure to be determined. The structure shows that the enzyme consists of two tight dimers that interact to form a homotetramer.[7]

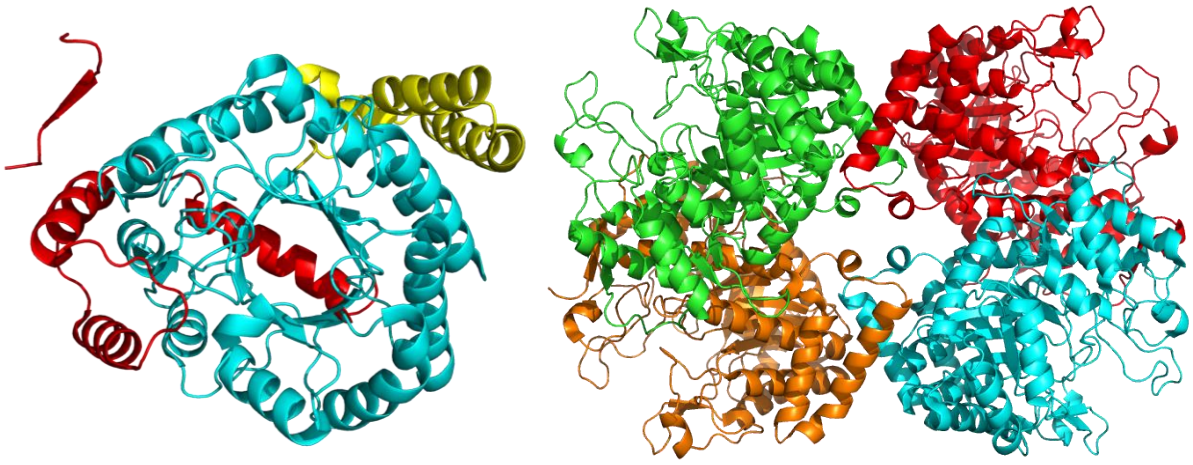


Figure 1.11: Monomeric (left) and tetrameric (right) structures of *Mtu*DAH7PS (PDB code 3NV8) showing the core $(\beta/\alpha)_8$ barrel (cyan), with the N-terminal extension (red) and internal extension (yellow) for the monomer and different colours representing different monomers for the tetramer.

As for the type I DAH7PS enzymes, this structure shows that each subunit consists of the same $(\beta/\alpha)_8$ barrel (Figure 1.11). Unlike type Iα and Iβ DAH7PS enzymes, *Mtu*DAH7PS contains two additional decorations to the core barrel that play a role in the allostery of the enzyme.[11,14,15] one extension consists of a β-strand followed by three helices extending from the N-terminus closing off the end of the barrel, with another extension consisting of two helices, extending the α2-β3 connecting loop.[46]

1.7 Active site of DAH7PS

Although the different types of DAH7PS enzymes (type Iα, Iβ, and II DAH7P synthases) have low overall sequence similarity, they all share the $(\beta/\alpha)_8$ barrel fold housing the catalytic machinery. Many of the residues that contribute to the active site are also common to all types of DAH7PS enzymes. The ordered sequential aldol condensation reaction involves the coupling of PEP and E4P to form DAH7P with a requirement for a divalent metal ion for the reaction to proceed. Three positively charged arginine (Arg) and two lysine (Lys) residues position the negatively charged PEP through hydrogen-bonding and salt bridge interactions.[47] These residues are conserved across all DAH7PS enzymes and in *Eco*DAH7PS, the residues involved in PEP binding are Arg92, Arg165, Arg234, Lys97 and Lys196 (Figure 1.12).[37]

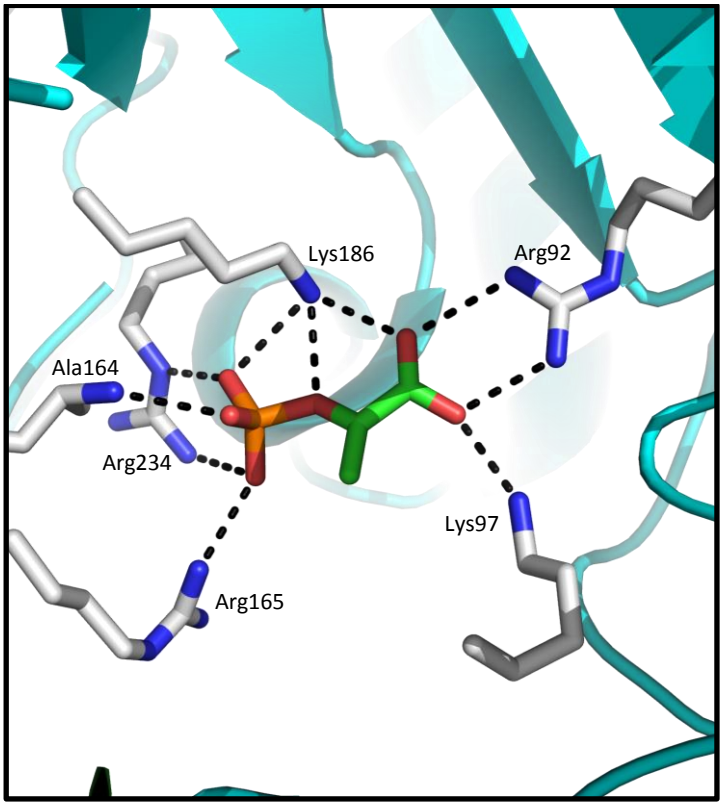


Figure 1.12: PEP binding site of *Eco*DAH7PS(Phe) (PDB code 1N8F) showing interaction (dashed lines) of residues (grey) with PEP (orange/green).

In the ordered sequential reaction carried out by DAH7PS, E4P is the second substrate to bind.[18] A crystal structure of *Tma*DAH7PS was solved with PEP and E4P bound.[41] Another crystal structure of *Sce*DAH7PS was obtained with PEP and glycerol 3-phosphate (G3P) bound.[47] These data gave a better understanding of the residues involved with the binding of E4P. E4P binding is mediated with the residues Lys, Pro, Arg (Thr/Ser), a conserved motif found in all DAH7PS enzymes. For *Eco*DAH7PS, the residues involved in E4P binding are Lys97, Pro98, Arg99 and Thr100.

In the active site, a divalent metal ion coordinates near the C-terminal end of the barrel.[37] An example of this is in *Eco*DAH7PS, where the divalent metal ion is coordinated in a distorted bi-pyramidal geometry to the residues Cys61, His268, Glu302 and Asp 326 (Figure 1.13).

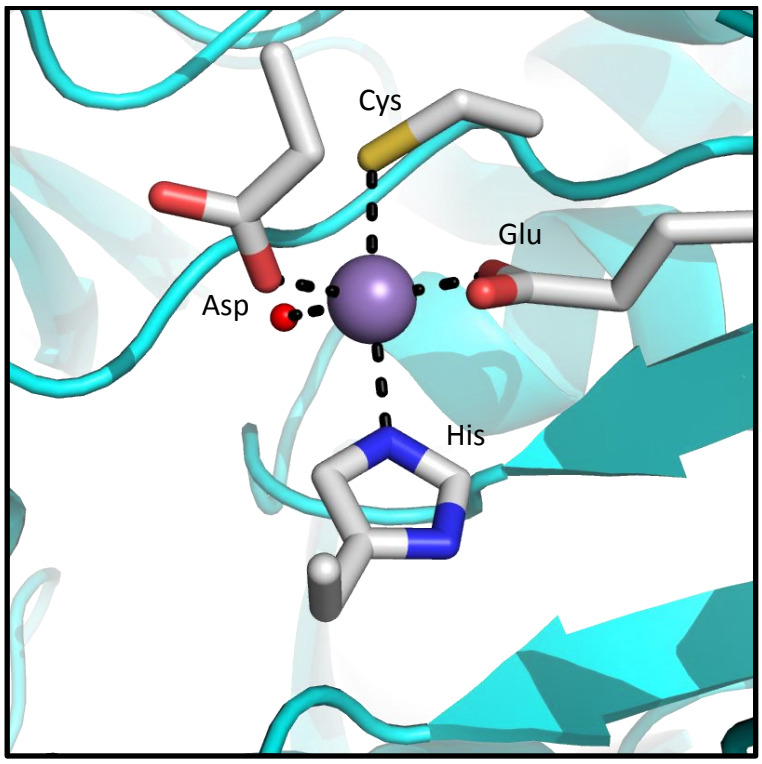


Figure 1.13: Metal binding site of *Eco*DAH7PS (PDB code 1N8F) showing water molecule (red) and metal ion (purple) coordinated to metal binding residues (grey).

1.8 DAH7PS regulation

Different methods of controlling the quantity of products formed are important to an organism as it ensures that biosynthesis of products only occurs as a response to metabolic demand. It is therefore in the best interest of the organism to control the rate at which products are being formed to prevent the wastage of resources. The main way that the shikimate pathway is controlled is through feedback regulation of DAH7PS by pathway intermediates or end products. For some types of DAH7PS enzymes, binding of these molecules remote to the active site leads to a conformational change that alters the active site, halting the enzyme from performing its intended function. Different species have employed different ways in how DAH7PS is regulated, with each type of enzyme appearing to have a unique method of allostery.[11,13,16]

All regulated DAH7PS enzymes from different organisms contain different types of decorations attached to the core $(\beta/\alpha)_8$ barrel. Obtaining crystal structures of DAH7PS enzymes has been very useful in determining the mechanism of allosteric inhibition

employed by the enzyme. A visual inspection of the decorations shows how they are folded, and can give some insight as to where allosteric inhibitors bind. Structural comparisons of ligand bound and non-ligand bound DAH7PS enzymes have also helped to determine the conformational changes that the enzyme undergoes upon ligand binding.[11,16]

Each of the three isozymes of *Eco*DAH7PS is feedback regulated by one of the three aromatic amino acids (Phe, Tyr and Trp) of the shikimate pathway. It has been shown that *Eco*DAH7PS(Phe) accounts for most of the DAH7PS activity (approximately 80 %) in *E. coli*, while *Eco*DAH7PS(Tyr) accounts for 20 %. The *Eco*DAH7PS(Trp) isozyme has the lowest contribution with less than 1 % of total DAH7PS activity.[48] *Salmonella typhimurium* and *Neurospora crassa* also produce three DAH7PS isozymes, whereas *S. cerevisiae* only produces two isozymes regulated by Phe and Tyr.[49]

A crystal structure of *Eco*DAH7PS(Phe) shows that the N-terminal extension and the internal extension located between the α5 and β6 strands of the $(\beta/\alpha)_8$ barrel create a single binding site for the inhibitor to bind to and inhibit the enzyme.[8,43] Comparisons of the ligand bound and non-ligand bound crystal structures of *Eco*DAH7PS(Phe) show that barely any conformational changes occur upon ligand binding (root-mean-square deviation (RMSD) of 0.7 Å between equivalent subunits), except for a slight disturbance in two active site loops involved in substrate binding (Figure 1.14). Two interconnected routes of conformational changes transmit from the allosteric Phe-binding site to the active site of DAH7PS. The transmitted conformational changes alter the ability of *Eco*DAH7PS to interact with its substrates, and is therefore no longer able to bind E4P, while only being able to bind PEP in a flipped conformation.

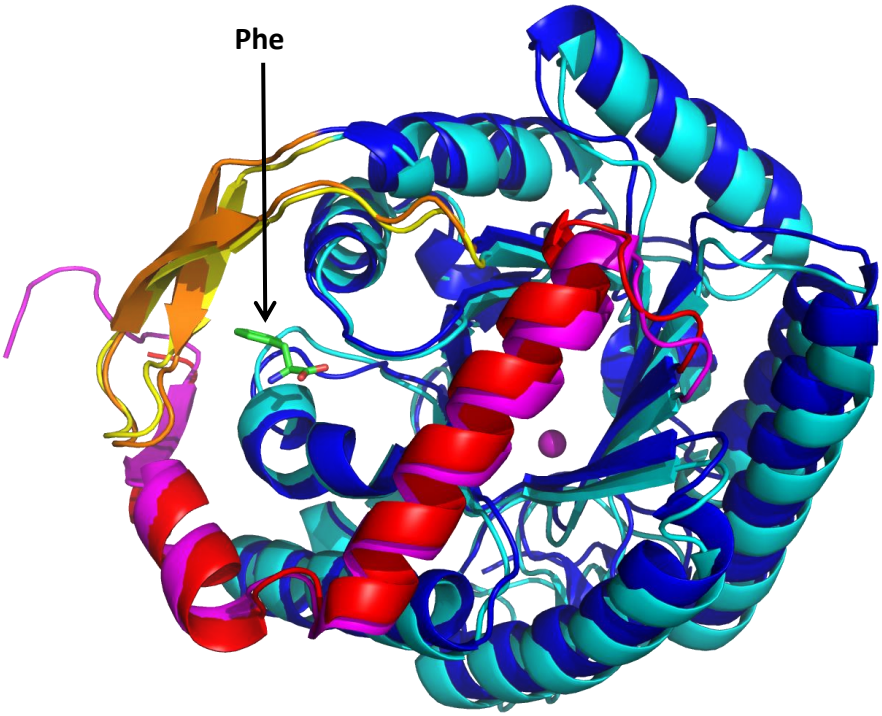


Figure 1.14: Superposition of a non-ligand bound (PDB code 1QR7) and ligand bound (PDB code 1KFL) crystal structure of *Eco*DAH7PS monomers. The non-ligand bound structure is shown in dark blue with the N-terminal and internal barrel extensions shown in red and orange respectively. The ligand bound structure is shown in cyan the N-terminal and internal barrel extensions shown in magenta and yellow respectively. The ligand Phe is shown as sticks coloured green with the metal ion shown as a purple sphere.

Some DAH7PS enzymes are simply unregulated. Belonging to the Iβ subfamily, *Pfu*DAH7PS and *Ape*DAH7PS have been structurally and functionally characterised, and have been found to consist of only the core $(\beta/\alpha)_8$ barrel with no added decorations (Figure 1.15).[14,15] Interestingly, both enzymes show no signs of regulation with none of the shikimate pathway products tested inhibiting the activity of the enzymes. This is unlike other Iβ subfamily DAH7PS enzymes that have evolved extensions to the core barrel that contribute to the regulation of the enzyme.

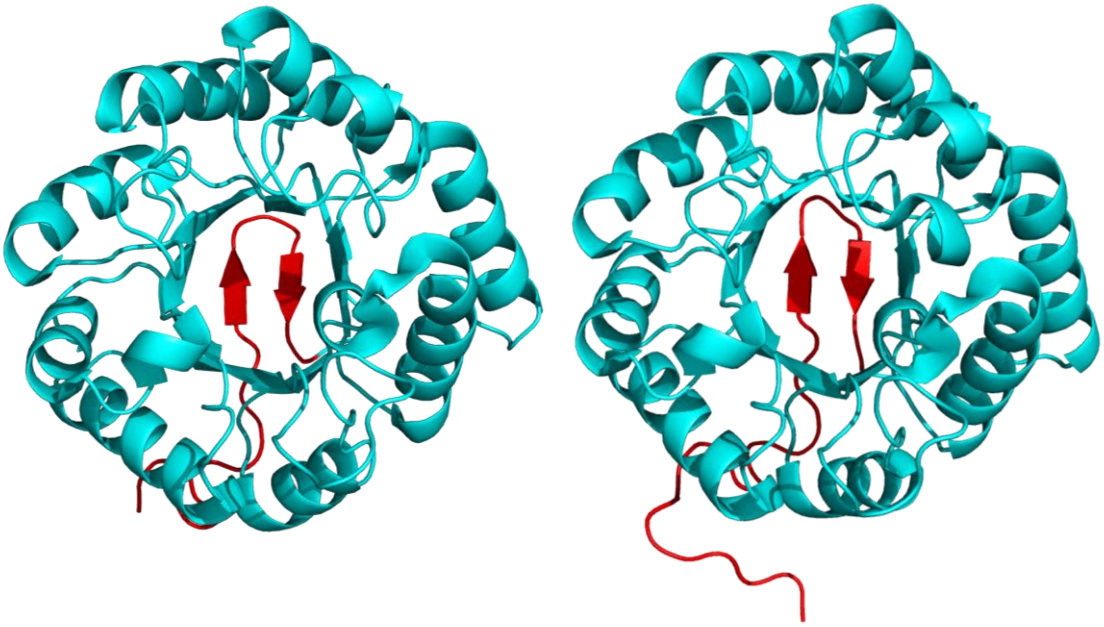


Figure 1.15: Monomeric structure of *Pfu*DAH7PS (left) (PDB code 1ZCO) and the monomeric structure of *Ape*DAH7PS (right) (PDB code 1VS1) showing the core $(\beta/\alpha)_8$ barrel (cyan) with the N-terminal extension (red).

*Tma*DAH7PS is strongly inhibited by Tyr and by Phe to a lesser extent.[38] Forming a truncated variant by the removal of N-terminal domain not only produced an enzyme that was catalytically more active, but one that was unaffected by Tyr or Phe. Furthermore, a crystal structure obtained of *Tma*DAH7PS in the presence of Tyr revealed that pairs of the regulatory domains from diagonally opposite subunits come together, forming a binding site for Tyr (Figure 1.16).[16] This reorganisation in the structure of *Tma*DAH7PS leads to the capping of the active site to restrict substrate binding, leading to enzyme inhibition. This showed that the FL domain was responsible for feedback regulation. More interestingly, a study carried out where the FL domain from *Tma*DAH7PS was fused to the unregulated *Pfu*DAH7PS, resulted in an enzyme that became regulated by Tyr and to a lesser extent by Phe.[50]

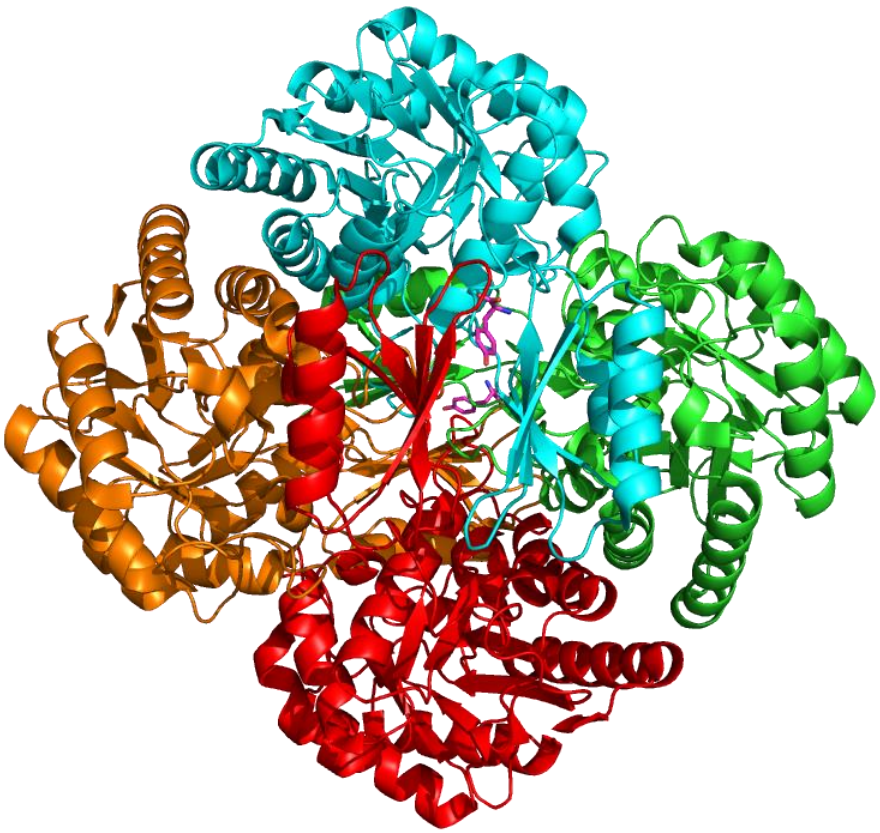


Figure 1.16: Crystal structure of *Tma*DAH7PS with Tyr bound (PDB code 3PG9). Different colours represent different monomers with Tyr ligands shown as sticks and coloured magenta.

The *M. tuberculosis* genome encodes for only one DAH7PS enzyme. This sole type II DAH7PS possesses a method of regulation much more complicated in comparison to other DAH7PS enzymes.[46] *Mtu*DAH7PS demonstrates little sensitivity to the three aromatic amino acids individually, yet significantly increased inhibition of the enzyme is observed in the presence of combinations of two or more different amino acids. These amino acids bind to the decorations that extend from the core $(\beta/\alpha)_8$ barrel. This synergistic type of inhibition allows for precise control over the quantity of end products being produced. More recently it was determined that while the combination of Phe and Trp caused significant inhibition of the enzyme, the addition of Tyr to the inhibitory couple led to almost complete loss of enzymatic activity.[11]

A ligand bound crystal structure of *Mtu*DAH7PS shows that there are three distinct allosteric binding sites for each aromatic amino acid.[11] Two of the allosteric sites are in the dimer and tetramer interfaces with the third binding site being located near the dimer interface on the

outside surface of the enzyme. Furthermore ligand bound and ligand free crystal structure comparisons of *Mtu*DAH7PS show no significant change in conformation of the enzyme (Figure 1.17).[46] Molecular dynamics simulation studies determined that the binding of ligands, caused a substantial increase in flexibility of the loop that forms part of the binding site for E4P.[17] This increased flexibility may mean that binding of the substrate E4P is entropically unfavourable, and may be the mechanism by which inhibition of the enzyme is achieved.

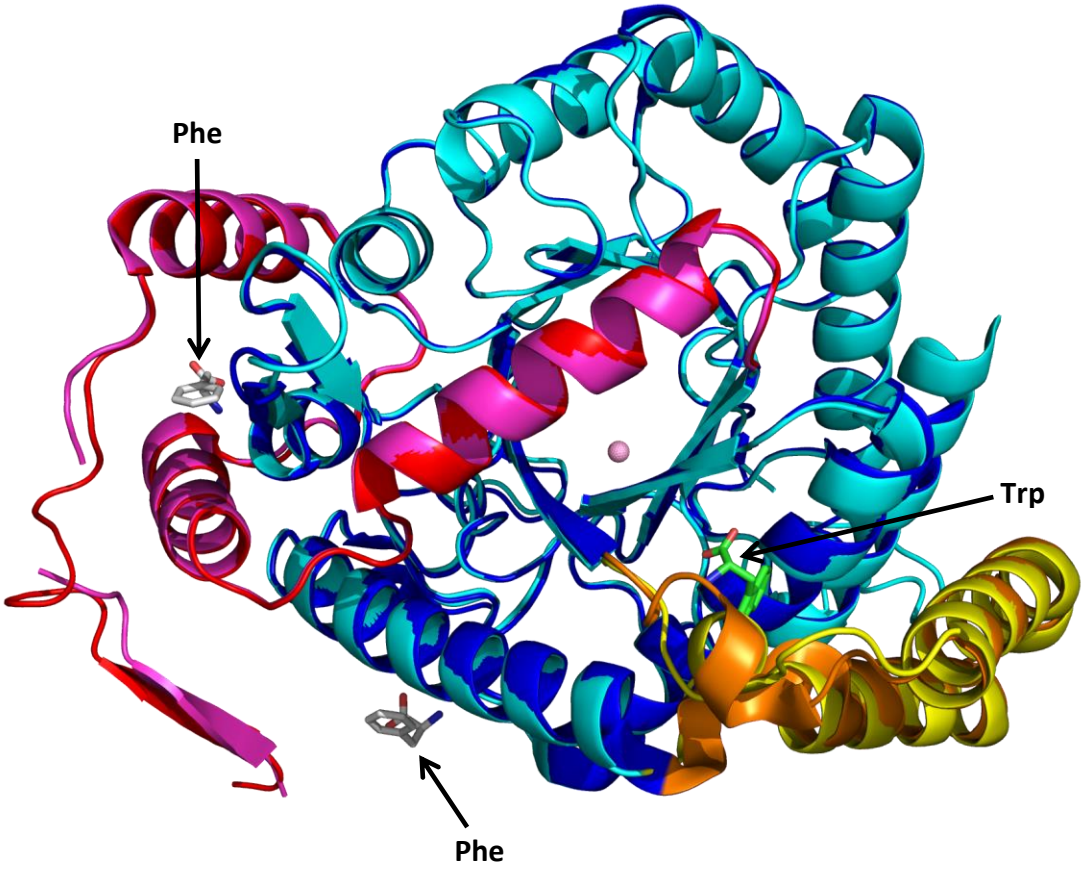


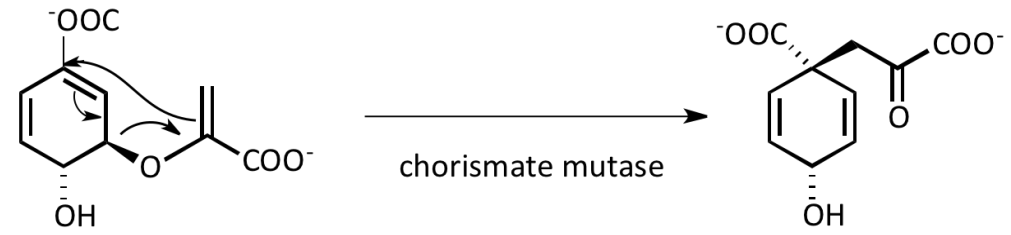
Figure 1.17: Superposition between a non-ligand bound (PDB code 2B70) and ligand bound (PDB code 3KGF) subunit of *Mtu*DAH7PS. The non-ligand bound structure is shown in dark blue with the N-terminal and internal barrel extensions shown in red and orange respectively. The ligand bound structure is shown in cyan the N-terminal and internal barrel extensions shown in magenta and yellow respectively. The ligands, Phe and Trp are shown as sticks coloured grey and green respectively with the metal ion shown as a purple sphere.

1.9 Chorismate mutase (CM)

Chorismate, the last compound produced in the main trunk of the shikimate pathway and at the branch point to a range of other aromatic metabolites, is the common precursor for vital

aromatic amino acid biosynthesis. After the shikimate pathway branches, one branch leads to the formation of anthranilate and thereon to Trp, whereas the other branch leads to the formation of prephenate, the precursor for both Tyr and Phe.[51]

Chorismate mutase (CM), first identified in 1965,[52] catalyses the conversion of chorismate to prephenate (Figure 1.18). Different types of CM enzyme structures can be found in different microorganisms, with most being grouped into the AroQ family and the rest being grouped into the AroH family. First observed *in E. coli,*[53] CM enzymes in the AroQ family are widespread and are mainly homodimeric, with each chain consisting of three α-helices connected by two turns (Figure 1.19).[54,55] CM enzymes in this family can be monofunctional or bifunctional. Being bifunctional means having the ability to catalyse two different reactions through two different catalytic sites found in a single protein chain. This occurs when CM is connected to another enzymatically active domain. Bifunctional AroQ CM domains were found to be fused to different enzymes such as prepehenate dehydratase, prephenate dehydrogenase or DAH7PS.[56,57]



chorismate **prephenate**

Figure 1.18: Conversion of chorismate to prephenate catalysed by chorismate mutase.

The other type of chorismate mutase enzymes are grouped into AroH family, characterised by a trimeric α/β barrel structure (Figure 1.19).[58,59] Chorismate mutase enzymes in this family are monofunctional as they are never fused to other functional domains like monofunctional AroQ enzymes, and therefore contain a single active site located at the monomer interfaces.

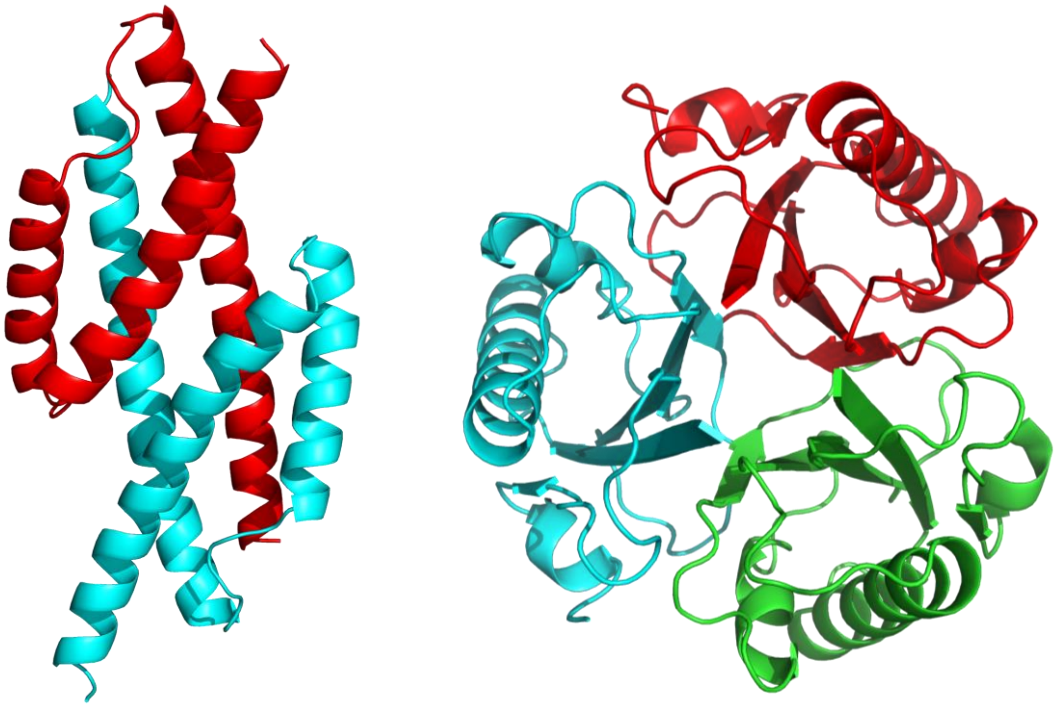


Figure 1.19: AroQ CM dimer structure (left) of *E. coli* (cyan and red) (PDB code 1ECM) and AroH CM trimer structure from *Bacillus subtilis* with (cyan, red, green) (PDB code 2CHT)

1.10 CM and DAH7PS fusion proteins

A slightly active CM domain covalently fused to the N-terminus of a subtype Iβ DAH7PS enzyme was first identified in *Bacillus subtilis* (*Bsu*DAH7PS).[58] The CM domain fused to the DAH7PS domain plays a role in regulating DAH7PS activity through feedback regulation. This occurs through binding of the branch point intermediates (chorismate and prephenate), to the CM domain, altering the DAH7PS domain active and leading to enzyme inhibition.[12,13,60]

Other characterised DAH7PS enzymes in different organisms, such as *Porphyromonas gingivalis* (*Pgi*DAH7PS),[12] and *Listeria monocytogenes* (*Lmo*DAH7PS),[61] also display fusion to an ancillary CM domain, with *Pgi*DAH7PS shown to be inhibited by prephenate and chorismate (Figure 1.20). It is thought that organisms evolved these fusion proteins for the sole purpose of regulation rather than to provide bifunctionality.[12] *Bsu*DAH7PS was found to exhibit both DAH7PS activity and extremely low CM activity.[13] It was determined however that *B. subtilis* contains another gene encoding for a second intracellular AroH CM domain responsible for the conversion of CA to PA.[60] This AroH CM is highly active, unlike the AroQ

CM which is redundant in activity by comparison and is really an ancillary to this other CM domain.

To provide further knowledge on the evolution of these fusion proteins, a study was carried out where the DAH7PS and CM domains of *Bsu*DAH7PS were separated by domain truncation.[13] The same study also examined the alternate fusion protein *Pgi*DAH7PS that also exhibited DAH7PS activity and low CM activity, in which the CM domain extension is fused C-terminal end of the catalytic barrel.[12] A domain truncation was also carried out on this enzyme and the results obtained were used for comparative studies. Results determined that the DAH7PS truncated variants from both enzymes only demonstrated DAH7PS activity. The CM truncated domain from *Bsu*DAH7PS only demonstrated CM activity, while the CM truncated domain from *Pgi*DAH7PS did not demonstrate DAH7PS or CM activity. Interestingly, whereas both wild-type enzymes were inhibited by prephenate and chorismate, the DAH7PS truncated mutants from both enzymes showed no signs of inhibition. In comparison to prephenate, chorismate was found to be a much weaker inhibitor for *Bsu*DAH7PS, with a sigmoid curve inhibition pattern in response to increasing concentrations of chorismate. It is believed that the weak inhibition pattern observed by chorismate is due to the weakly active CM domain converting the substrate to prephenate, which in turn acts as the true inhibitor of *Bsu*DAH7PS. Intracellularly, chorismate would be converted to prephenate at a much faster rate due to the other AroH CM protein that is highly active in *B. subtilis*, leading to the inhibition of the enzyme. The results obtained presented compelling evidence that these organisms evolved these fusion proteins for the sole purpose of regulation instead of bifunctionality.[13]

*Lmo*DAH7PS is another fusion protein that consists of the core $(\beta/\alpha)_8$ barrel most similar to the catalytic barrel domain of *Tma*DAH7PS, with 55 % sequence identity.[41,61] The core barrel is fused to a CM domain through a linker region at the N-terminus. The crystal structure of the enzyme shows that it forms a tetramer with two CM domains emerging from diagonally opposite sides of the tetramer and to form a CM dimer (Figure 1.20). The CM dimers on each side of the tetramer are believed to stabilise the tetrameric state of the enzyme in the same way the FL domains stabilise *Tma*DAH7PS. It is thought that the CM dimers are loosely associated with the DAH7PS tetramer where binding of prephenate or chorismate to the

regulatory domain would lead to a conformational change. This proposed conformational change may cause catalytic and regulatory domain interactions, either blocking the active site, or inducing conformational changes in the catalytic domain.[61]

Another type of interaction between CM and DAH7PS was discovered for *M. tuberculosis*, which differed from the others in that it is a non-covalent interaction between independent CM and DAH7PS proteins.[62] The genome of *M. tuberculosis* encodes for a weakly active chorismate mutase domain (*Mtu*CM). This weak enzymatic activity was accounted for by the absence of some of the conserved active site residues present in the AroQ family. It was determined that upon complex formation between *Mtu*CM and *Mtu*DAH7PS, *Mtu*CM became much more active with a more than a 100 fold increase in catalytic activity, whereas *Mtu*DAH7PS activity was unaffected. In the presence of Phe, Tyr or Trp, *Mtu*CM showed no signs of inhibition. Upon complex formation with *Mtu*DAH7PS, Phe and Tyr inhibited the enzyme by 70 % and 40 % respectively, with synergistic inhibition occurring In the presence of both aromatic amino acids, almost completely abolishing CM enzymatic activity.[63]

A crystal structure of the protein complex showed that CM dimers associated with opposite ends of the *Mtu*DAH7PS tetramer (Figure 1.20). Inspection of the crystal structure revealed that *Mtu*DAH7PS does not supply residues to the active site of *Mtu*CM to increase activity, however it causes a conformational change in *Mtu*CM that leads to a displacement of residues in the active site for more favourable catalysis.[63]

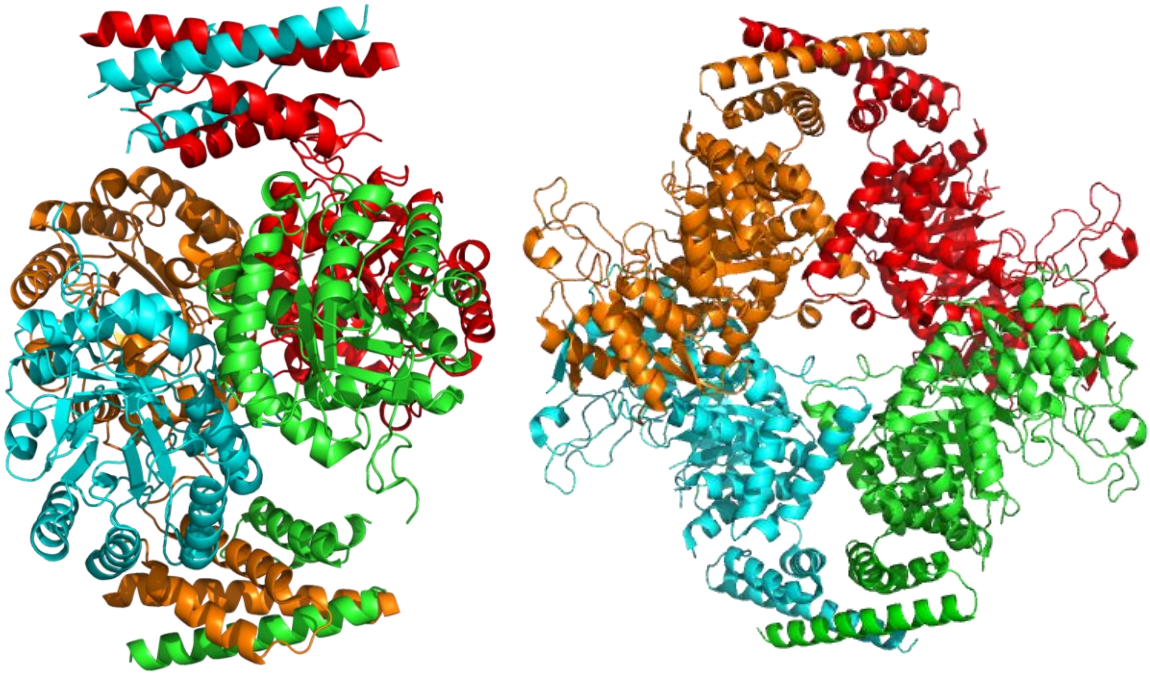


Figure 1.20: Tetramer structure of *Lmo*DAH7PS (PDB code 3TFC) (left), and *Mtu*DAH7PS-CM (PDB 2W19) (right) complex with different colours representing different monomers, each fused to an N-terminal CM domain through a linker region.

1.10.1 Geobacillus sp DAH7PS

Geobacillus sp is a Gram negative bacterium which is a thermophile, and thrives in a variety of locations at temperatures as high as 75 °C.[64] Sequence analysis, along with a crystal structure of *Geobacillus sp* DAH7PS (*Gsp*DAH7PS^{WT}),[65] shows that this enzyme also has fused to the CM domain at the N-terminus (Figure 1.21).

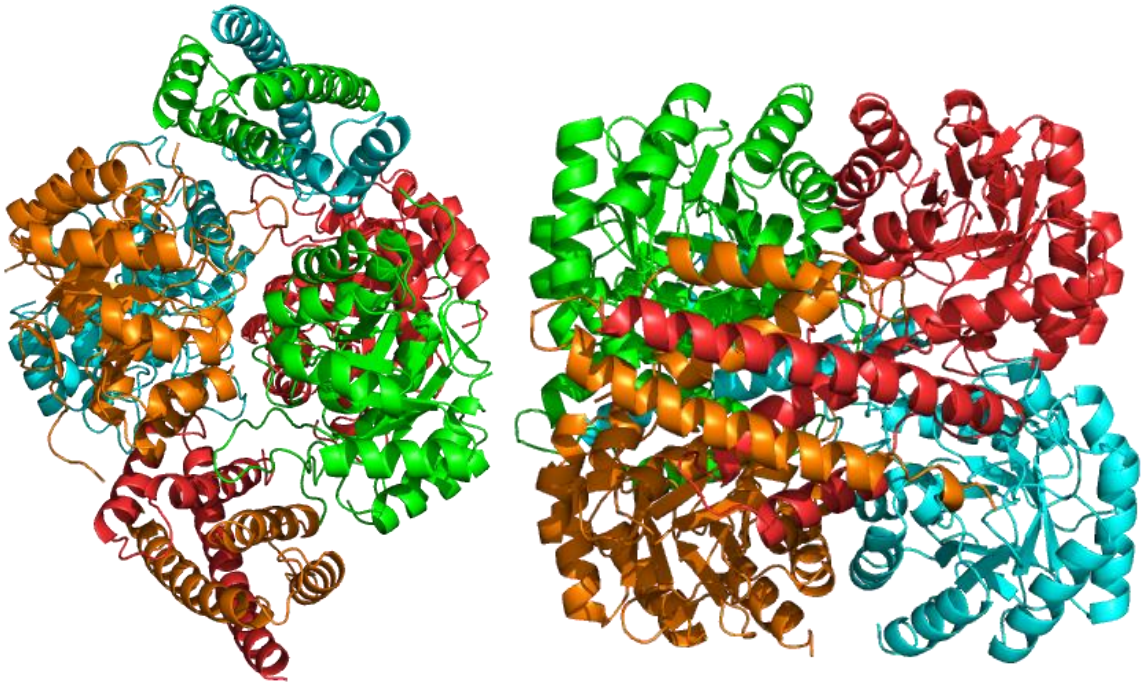


Figure 1.21: Tetramer structure of *Gsp*DAH7PS^{WT} at different angles, with different colours representing different monomers, each fused to an N-terminal CM domain through a linker region.

*Gsp*DAH7PS^{WT} crystallises as a tetramer, where the quaternary structure is held together by the CM domains from diagonally opposite sides of the tetramer interlocking with one another, in a similar way to the *Lmo*DAH7PS structure. Each subunit consists of the core $(\beta/\alpha)_8$ barrel bound to a CM domain on the N-terminus through a flexible linker (Figure 1.22). This CM domain is thought to play a role in the regulation of the enzyme.

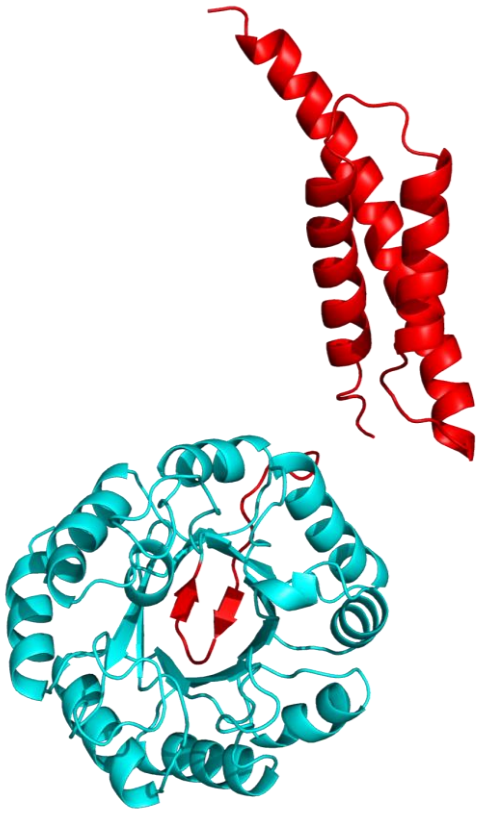


Figure 1.22: Monomeric structure of *Gsp*DAH7PS^{WT} showing the core $(\beta/\alpha)_8$ barrel (cyan) with the N-terminal CM regulatory domain (red), connected through a linker region (not shown due to lack of electron density).

1.11 Objectives of thesis

The aim of this project was to contribute towards the research previously carried out on DAH7PS enzymes, in order to determine how these enzymes evolved such diverse methods of regulation. The idea that these enzymes have recruited a variety of regulatory machinery, by attaching small regulatory domains to the core barrel for the sole purpose of feedback regulation was tested. This hypothesis was examined by carrying out a domain truncation of the type Iβ *Gsp*DAH7PS^{WT} at the linker region, in order to obtain a catalytic domain lacking the N-terminal domain (*Gsp*DAH7PS^{DAH7PS}), and a regulatory domain without the catalytic domain (*Gsp*DAH7PS^{CM}). After separating the catalytic and regulatory domains, both variants were fully characterised, and information obtained from each domain was compared to the respective catalytic and regulatory domains of the wild-type enzyme, which was also characterised in studies described in this thesis.

This study aimed to:

- Determine if the regulatory CM domain does in fact contribute to allosteric regulation of the enzyme, and if so, help predict the likely mechanism of inhibition.
- Investigate if the CM domain plays a role in the quaternary structure of the wild-type enzyme and what effects the removal of this catalytic domain has.
- Gain more insights to the different modes of regulation of the DAH7PS enzymes.

Chapter 2

Purification and biochemical characterisation of type Iβ *Geobacillus sp* DAH7PS

2.1 Overview

The genome of the thermophile *Geobacillus sp*, encodes for a type Iβ DAH7PS enzyme consisting of a catalytic domain fused to a regulatory CM domain, as determined by sequence analysis and the structure of the enzyme.[65] The aim of the work described in this chapter was to fully characterise the wild-type enzyme and examine in detail both the independent regulatory and catalytic domains. Data obtained has been used for comparative studies with previously characterised DAH7PS enzymes, particularly to other CM fused DAH7PS enzymes. The information gained by this study provides a greater understanding of how these enzymes evolved to incorporate the regulatory decorations to the core barrel.

The initial characterisation of *Gsp*DAH7PS^{WT} begun by the expression and purification of the enzyme. Differential scanning fluorimetry (DSF), metal activation, temperature dependency and kinetic study experiments were then carried out. DSF determined the stabilising effects different additives had on *Gsp*DAH7PS^{WT}. Metal activation and temperature dependency studies were conducted to determine the most activating metal ion for the enzyme, and the temperature for highest enzymatic activity. Kinetic studies involved obtaining Michaelis-Menten constants and carrying out an inhibition study to determine if the enzyme is inhibited. Analytical ultracentrifugation was used to determine the quaternary structure of the enzyme in solution.

2.2 Expression and purification

The purification protocol for *Gsp*DAH7PS^{WT} (with a C-terminal non cleavable hexaHis-tag ligated into the PET-28b (+) vector) was developed with reference to its predicted heat stability and calculated isoelectric point (pI) of 5.99. *E. coli* cells containing the plasmid encoding for *Gsp*DAH7PS^{WT} were grown, induced with isopropyl-1-thiol-β-D-galactopyranoside (IPTG), harvested and lysed. The lysed cells were heat treated at 60 °C for 30 minutes and left to cool to room temperature in an attempt to remove other contaminating proteins. Since *Gsp*DAH7PS^{WT} is naturally expressed in a thermophilic host,[64] it was expected to able to withstand heat treatment, while other proteins originating from *E. coli* host cells would denature. The heat-treated supernatant was centrifuged to remove cell debris and precipitated protein. The resulting supernatant required further purification steps to remove remaining contaminants (Figure 2.1).

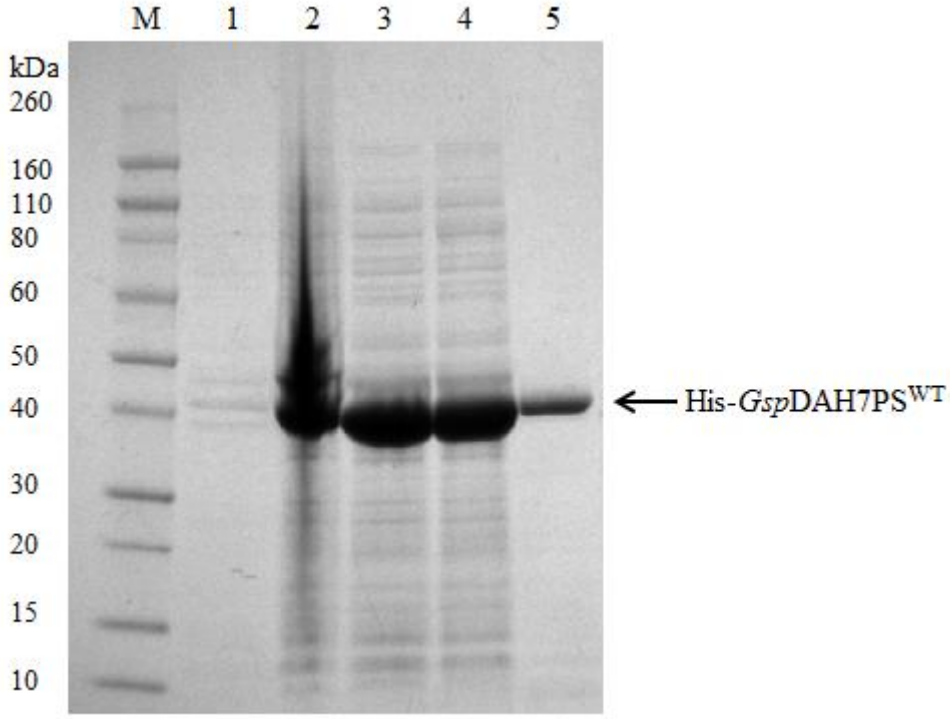


Figure 2.1: SDS-PAGE analysis for *Gsp*DAH7PS^{WT} before induction (lane 1), after induction (lane 2), whole cells containing *Gsp*DAH7PS^{WT} (lane 3) after lysis and heat treatment (lane 4), and in the supernatant after centrifugation (lane 5).

Since *Gsp*DAH7PS^{WT} contains a C-terminal hexaHis-tag, a nickel-based affinity column was used in the next purification step. The supernatant was loaded onto the column in buffer containing a low imidazole concentration (20 mM) and eluted using a linear gradient of

increasing imidazole concentration. The eluted fractions containing *Gsp*DAH7PS^{WT} (as determined by SDS-PAGE) were collected, concentrated, and then the protein was further purified using size exclusion chromatography (SEC). The protein eluted from the SEC column in a single peak with a shoulder to the right of the peak (Figure 2.2).

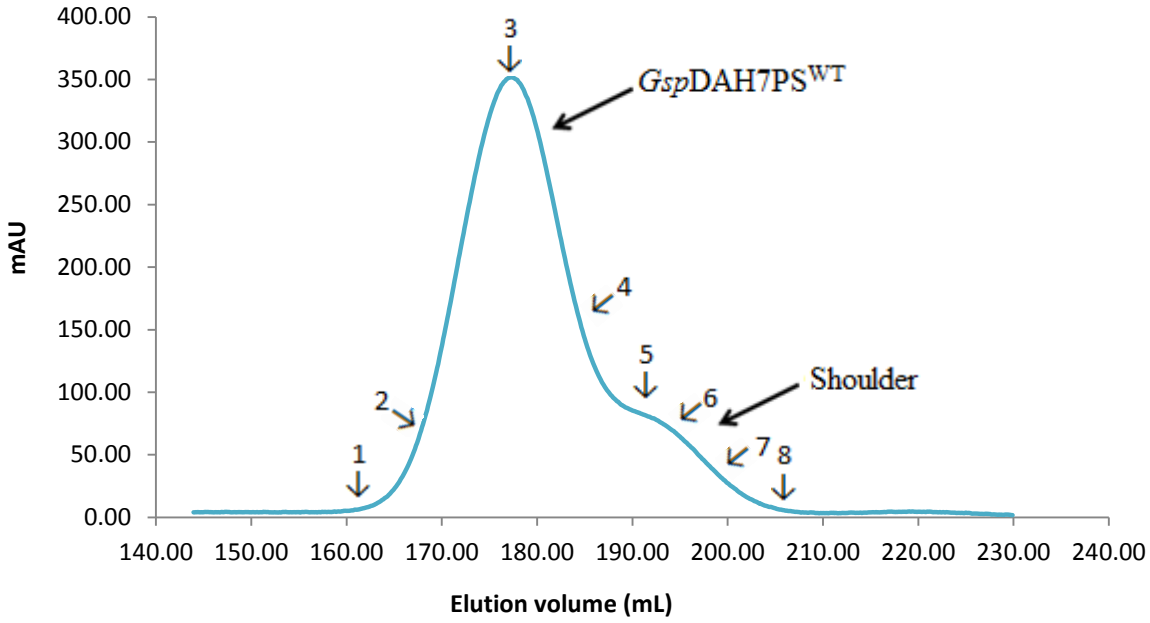


Figure 2.2: SEC trace of eluted *Gsp*DAH7PS^{WT} in SEC buffer (10 mM BTP, pH 7.4, 40 mM KCl, 200 µM PEP and 10 µM EDTA) showing the protein peak with a shoulder to the left of the peak. Numbers 1 to 8 represent the lane number on the SDS-PAGE gel of where the fractions from the SEC run were taken (Figure 2.3).

SDS-PAGE analysis determined that both the peak and its shoulder represented a protein with a molecular mass corresponding to *Gsp*DAH7PS^{WT} (Figure 2.3). The peak and the shoulder both contained DAH7PS activity as determined by enzyme activity assays, suggesting *Gsp*DAH7PS^{WT} had been purified. Fractions corresponding to the larger peak only were collected, concentrated, divided into aliquots, flash frozen using liquid nitrogen, and stored at -80 ˚C.

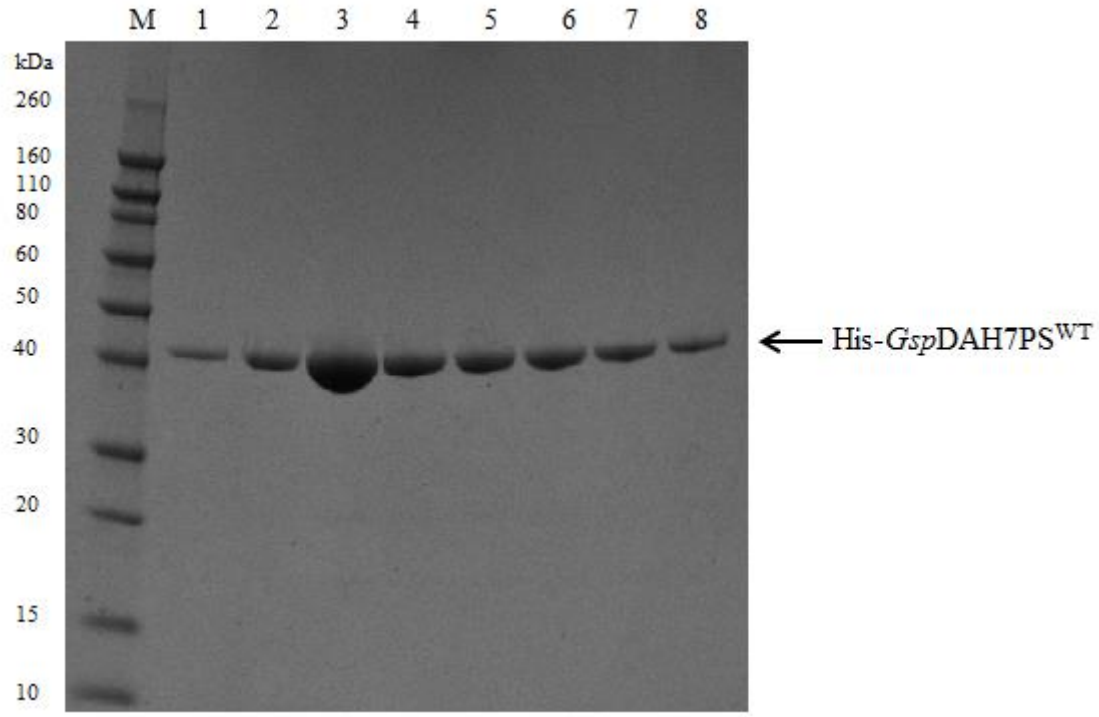


Figure 2.3: SDS-PAGE gel of final collected fractions after SEC of *Gsp*DAH7PS^{WT} Lanes 1-4 showing fractions collected from the peak (Figure 2.2). Lanes 5-8 showing fractions collected from the shoulder.

2.3 Mass spectrometry

The molecular mass of *Gsp*DAH7PS^{WT} was determined using electrospray ionisation mass spectrometry to ensure the purified enzyme was of the expected mass calculated from the amino acid sequence using ProtPram. The enzyme had a mass of 41,373 Da, which is in close agreement with the calculated mass of 41,374 Da.

2.4 Differential scanning fluorimetry

To investigate the heat stability of *Gsp*DAH7PS^{WT} and to probe any changes to the thermal stability of this protein under different conditions (no metal-control, EDTA, Co^{2+}, Cd^{2+}, Zn^{2+}, Mg^{2+}, Ni^{2+}, Pb^{2+}, Ca^{2+}, chorismate and prephenate). DSF experiments were carried out to determine the protein denaturation temperatures (T_m) (Figure 2.4). This method utilises a dye that fluoresces in the presence of hydrophobic amino acids. As the temperature

increases, proteins begin to denature, exposing hidden hydrophobic residues, which interact with the dye and lead to an increase in fluorescence.[66] The temperature at which a protein denatures is therefore determined as the point with the highest rate of increase in fluorescence.

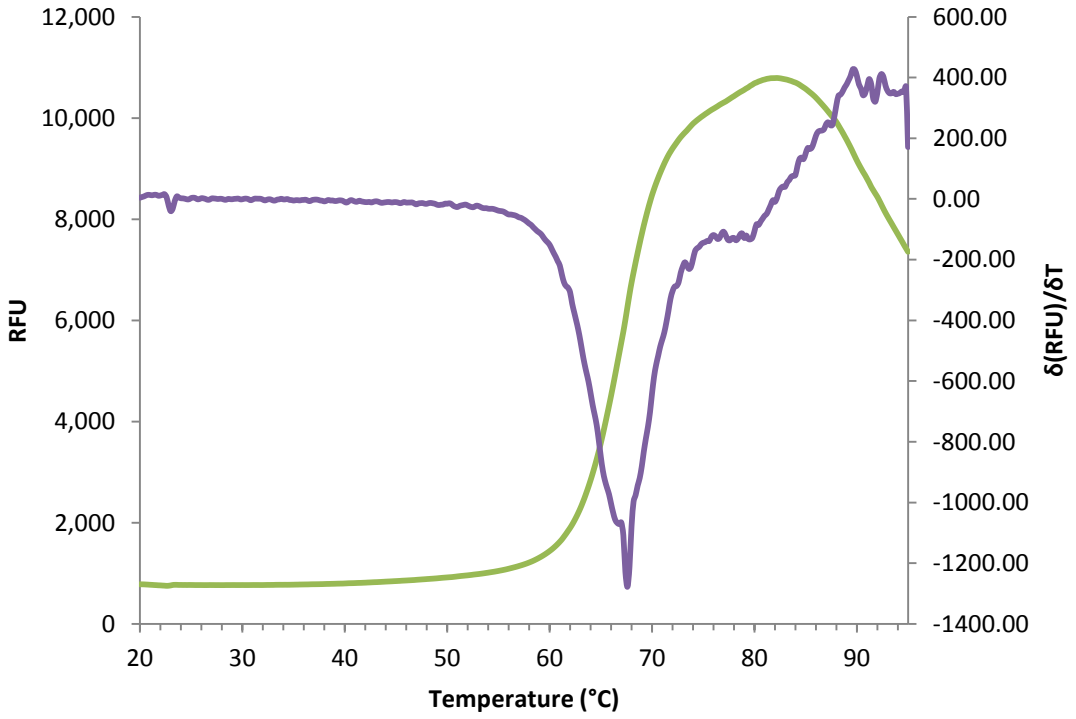


Figure 2.4: DSF trace of *Gsp*DAH7PS^{WT} in the presence of 100 µM EDTA, where RFU is the relative fluorescence units. Shown in purple is the negative of the first derivative on the secondary axis. T_m is determined as the steepest part of the RFU trace and the minimum on the negative derivative.

The T_m of the wild-type enzyme with no metal added (control) was determined to be 76.7 ± 0.2 °C. This value is very similar to the other T_m values obtained for the enzyme in the presence of a range of different divalent metal ions. A significantly lower T_m of 67.8 ± 0.3 °C in the presence of 100 µM of EDTA was recorded. This difference in T_m in the presence of EDTA, compared to other conditions, suggests that there are traces of metal present in the protein preparation that are having a stabilising effect on the enzyme. This is further supported by the metal activation study on *Gsp*DAH7PS^{WT}, which showed that enzyme was still active without metal additives and almost completely deactivated in the presence of EDTA, suggesting metal presence in the protein preparation (section 2.5). This would explain

the T_m value obtained for the control, as EDTA is a chelator that removes metal ions from the enzyme. The presence of the metals Cd^{2+} and Zn^{2+} causes two distinct rises in florescence, which are consistent with the wild-type enzyme having a two stage denaturation response to the increase in temperature, (Figure 2.5) with initial T_m values of 53.1 ± 0.3 °C and 42.5 ± 0.4 °C respectively and secondary T_m values of 78.4 ± 0.7 °C and 81.9 ± 0.7 °C respectively. The initial T_m values suggest that both metals are destabilising, while the secondary T_m values suggest that they are stabilising to the enzyme in comparison to the T_m value obtained in the presence of EDTA (Figure 2.6). The two-stage denaturation response appears to occur with *Gsp*DAH7PS^{WT} in the presence of other metal ion additives as well, however, the peak for the second rise in fluorescence seems to be masked by the first peak (Figure 2.4).

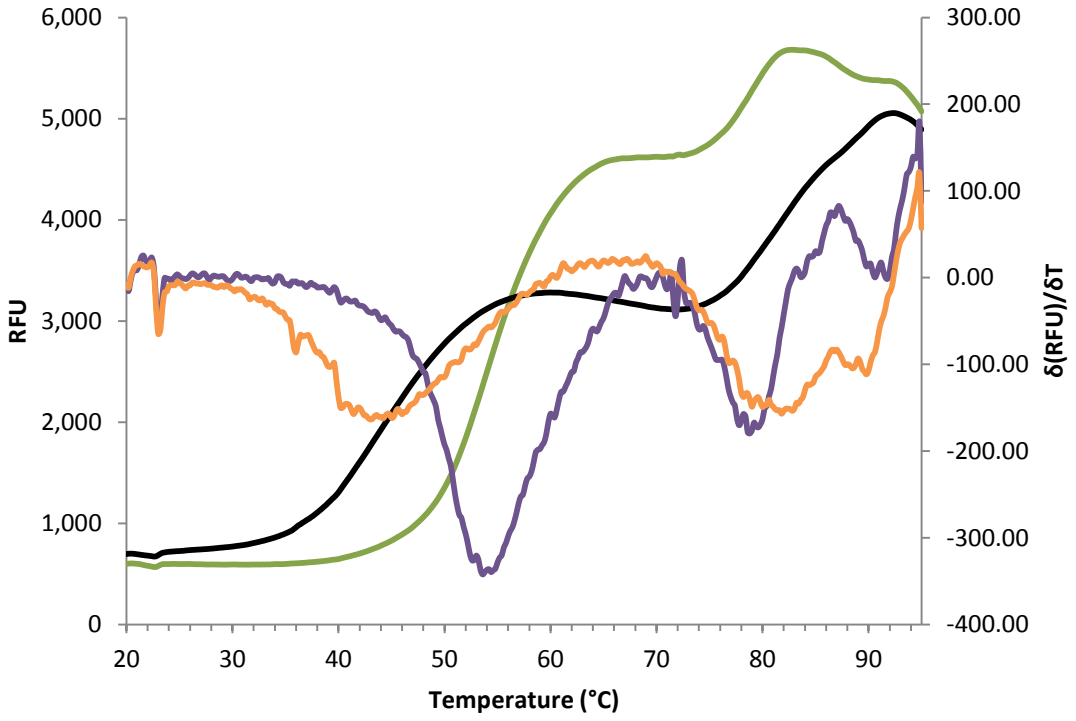


Figure 2.5: DSF trace of *Gsp*DAH7PS^{WT} in the presence of 100 µM Cd^{2+} (purple and green lines) and 100 µM Zn^{2+} where RFU is the relative fluorescence units. The negative of the first derivative on the secondary axis is shown in purple for Cd^{2+} and in orange for Zn^{2+}. T_m is determined as the steepest part of the RFU trace and the minimum on the negative derivative.

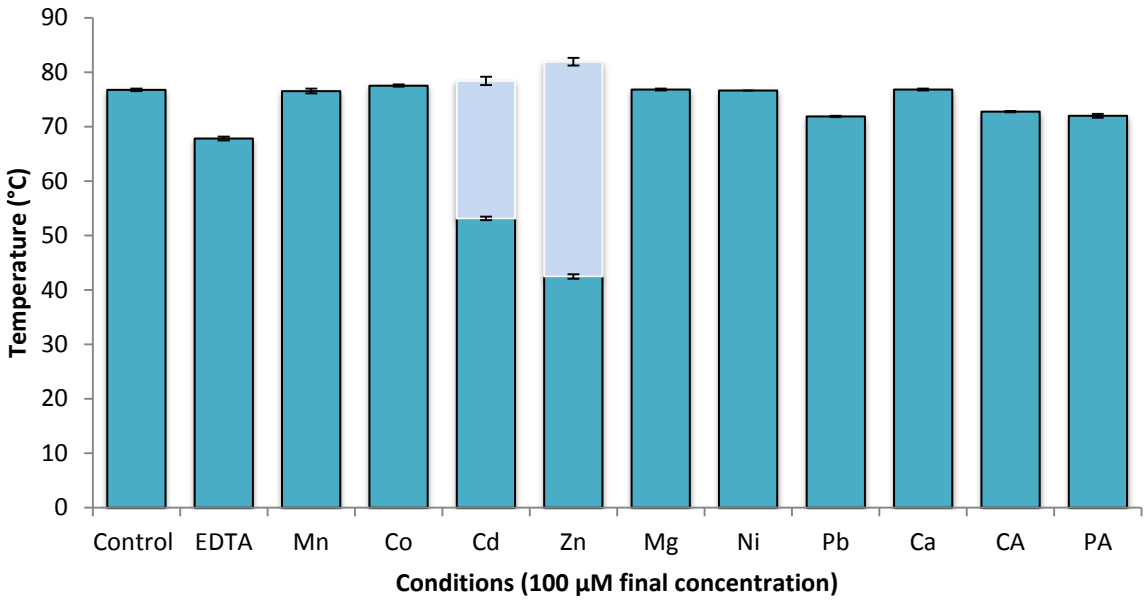


Figure 2.6: DSF data showing T_m of *Gsp*DAH7PS^{WT} in the presence of different conditions (no metal-control, EDTA, Co^{2+}, Cd^{2+}, Zn^{2+}, Mg^{2+}, Ni^{2+}, Pb^{2+}, Ca^{2+}, CA and PA at 100 µM with light blue bars indicating the secondary T_m for the enzyme.

All of the conditions with added metal ions had a stabilising effect on the wild-type enzyme, with T_m values higher than that in the presence of EDTA. Pb^{2+} was the least stabilising metal ion to the wild-type enzyme with a T_m of 71.9 ± 0.1 °C (Figure 2.6). In the conditions with 100 µM PA or CA present, there was a slight destabilising effect on *Gsp*DAH7PS^{WT} in comparison to the no metal added (control) condition with a T_m of 72.7 ± 0.1 °C and 72 ± 0.3 °C respectively.

2.5 Metal activation

All DAH7PS enzymes characterised to date have been shown to be dependent on a divalent metal ion for activity.[67] A metal activation study was carried out in order to determine if the activity of *Gsp*DAH7PS^{WT} is metal dependent, and if so, which metal is most activating for enzymatic activity. Buffer and substrate solutions used in metal activation studies were treated with Chelex® (an ion-exchange resin that removes polyvalent metal ions). A range of metal ions was used, along with metal chelators for the study. Cd^{2+} was found to be the most activating by a significant margin, whereas Zn^{2+} was least activating (Figure 2.7).

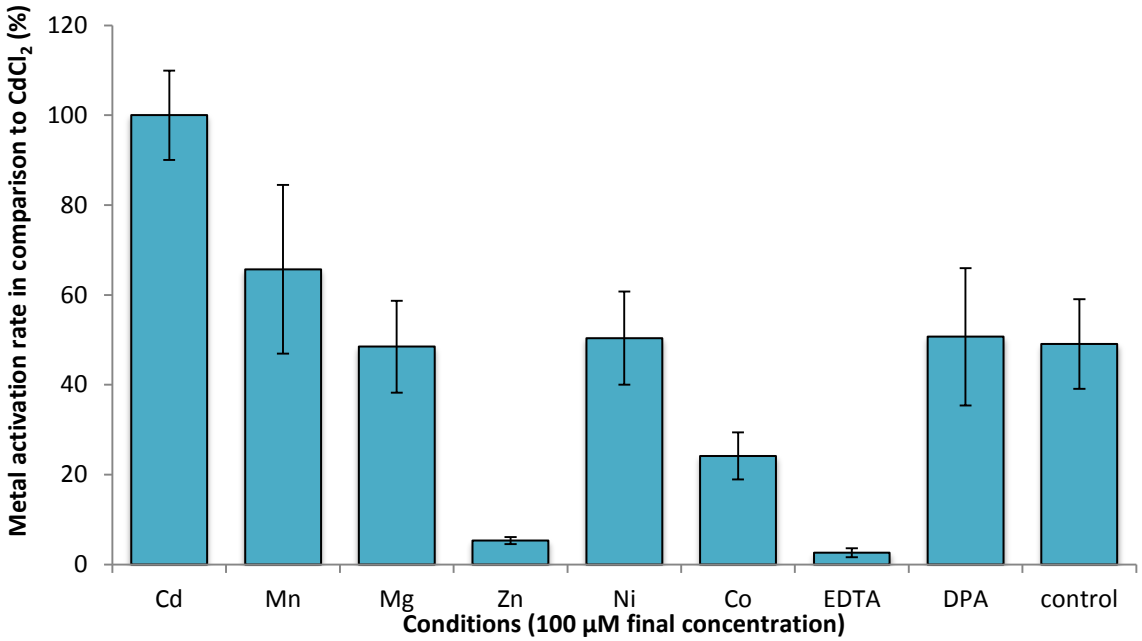


Figure 2.7: Metal activation of *Gsp*DAH7PS^{WT} representing the different activation abilities of different metals in comparison to Cd^{2+} at a concentration of 100 µM. Also shown is the degree of inactivation of the wild-type enzyme in the presence of the metal chelators EDTA and DPA at a concentration of 100 µM. Reactions were carried out at 60 °C, with 195 µM PEP and 197 µM E4P in 50 mM BTP buffer (pH 6.8).

It was demonstrated that the enzyme was still active without any additives (control), suggesting metal presence in the protein preparation, consistent with the observations of the DSF study (section 2.4). The metal chelator dipicolinic acid (DPA) at 100 µM did not deactivate *Gsp*DAH7PS^{WT} as the enzyme had very similar activity to the control condition (no additives). In the presence of chelator EDTA, at a concentration of 100 µM, the enzyme only demonstrated 3 % of the activity that was observed in the presence of Cd^{2+}.

This study confirms that *Gsp*DAH7PS^{WT} is dependent on a metal ion for activation like all other characterised DAH7PS enzymes. The enzyme demonstrated very little activity in the presence of EDTA with a clear preference for Cd^{2+}, which was most activating of the metal ions tested. While there has been no connection made with the type of DAH7PS enzyme and their metal preference, it is interesting to note that another type Iβ DAH7PS enzyme

fused to a CM domain isolated from the *Bacillus* genus, *Bsu*DAH7PS, is also most activated by Cd^{2+}.[13]

2.6 Temperature dependency study

Geobacillus sp is a thermophile that thrives at temperatures as high as 70 °C.[64] To determine the optimal temperature for *Gsp*DAH7PS^{WT} activity, a temperature dependency study was carried out using a continuous spectrophotometric assay following the loss of PEP at 232 nm.[40] Substrates (PEP and E4P), metal (Cd^{2+}), buffer (50 mM BTP, pH adjusted to the required temperature) and enzyme concentrations were kept constant while varying the temperature of the reaction mixture. Once the reaction mixture in the cuvette had equilibrated to the required temperature, enzyme was added and incubated for one minute before initiation of the reaction with E4P. The experiment was later repeated with the enzyme incubated for longer period of five minutes, to determine the effects of a longer incubation time (Figure 2.8). It has been determined that the substrate E4P is unstable at elevated temperatures, with less than half of the E4P concentration (64 µM) remaining after 30 minutes incubation at 60 °C.[40] Therefore it was important to initiate the reactions with E4P to minimise substrate loss, especially for the assays carried out at temperatures exceeding 60 °C.

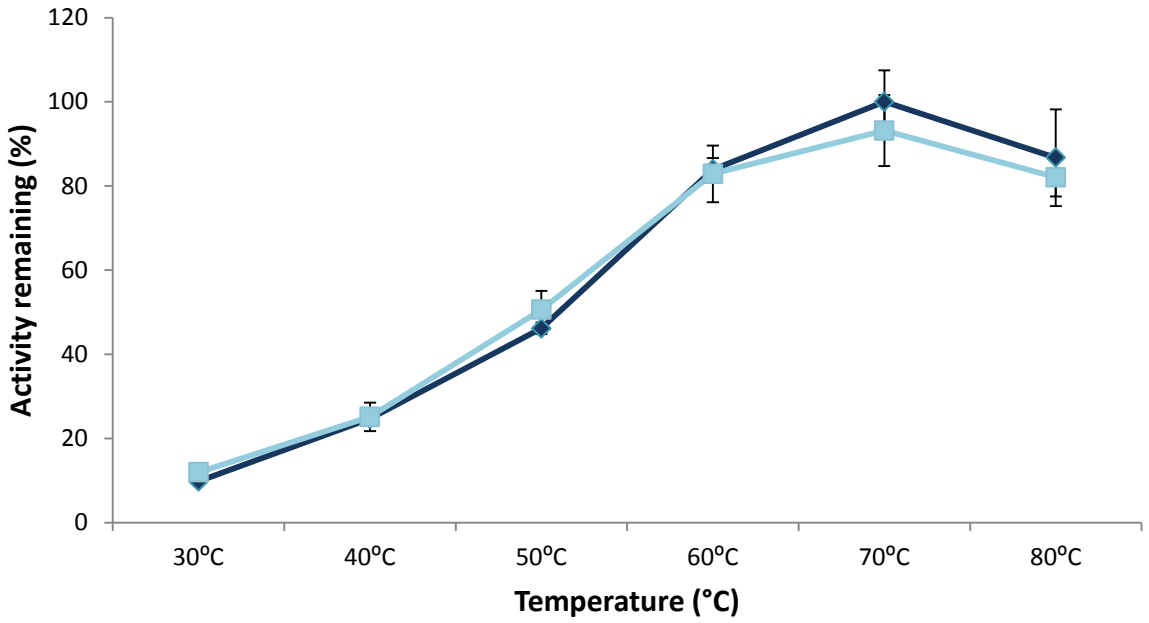


Figure 2.8: Temperature dependency of *Gsp*DAH7PS^{WT} showing enzyme activity at different temperatures. Substrate (325 µM PEP, 475 µM E4P), metal (100 µM Cd^{2+}), buffer (50mM BTP, pH 6.8) and enzyme (5 µL at 0.2 mg/mL) concentrations were kept constant. Enzyme was incubated at desired temperature with the reaction mixture for either one minute (shown in dark blue) or five minutes (shown in light blue)

*Gsp*DAH7PS^{WT} is most active at 70 °C and least active at 30 °C under the given conditions. Activity increases as temperature increases, up to 70 °C, and then drops slightly at 80 °C. This result can be explained in the context of the DSF data for the wild-type enzyme, which had a T_m value of 78.4 ± 0.7 in the presence of Cd^{2+} (section 2.4). At 80 °C, the enzyme is likely to be denaturing, and a reduction in enzymatic activity is observed. While *Gsp*DAH7PS^{WT} did have an initial T_m of 53.1 ± 0.3 °C in the presence of Cd^{2+}, according to the DSF data, enzyme activity did not appear to be affected at 60 °C. This would suggest that at 60 °C, the catalytic domain of the wild-type enzyme is still fully functional and not denaturing as no decrease in DAH7PS activity is observed. Incubation of the enzyme with the reaction mixture for either one or five minutes did not appear to have an effect on the activity of the enzyme, although the activities at the elevated temperatures were slightly lower, consistent with the denaturation of the enzyme (Figure 2.8).

2.7 Temperature dependency study (chorismate mutase activity)

Preliminary activity assays determined that *Gsp*DAH7PS^{WT} also had CM activity, consistent with the prediction of a CM fusion protein based on sequence analysis. Another temperature dependency study was carried out on *Gsp*DAH7PS^{WT} for CM activity. A continuous spectrophotometric assay following the loss of CA at 274 nm was used, where the substrate CA, buffer (pH adjusted for experimental temperatures) and enzyme concentrations were kept constant, while only varying the temperature (30 °C, 40 °C, 50 °C, 60 °C, 70 °C and 80 °C). Once the reaction mixture in the cuvette had equilibrated to the required temperature, enzyme was added to the mixture and incubated for one minute before initiation of the reaction with CA. The experiment was later repeated with the enzyme incubated for longer period of five minutes, to determine the effects of a longer incubation time (Figure 2.9).

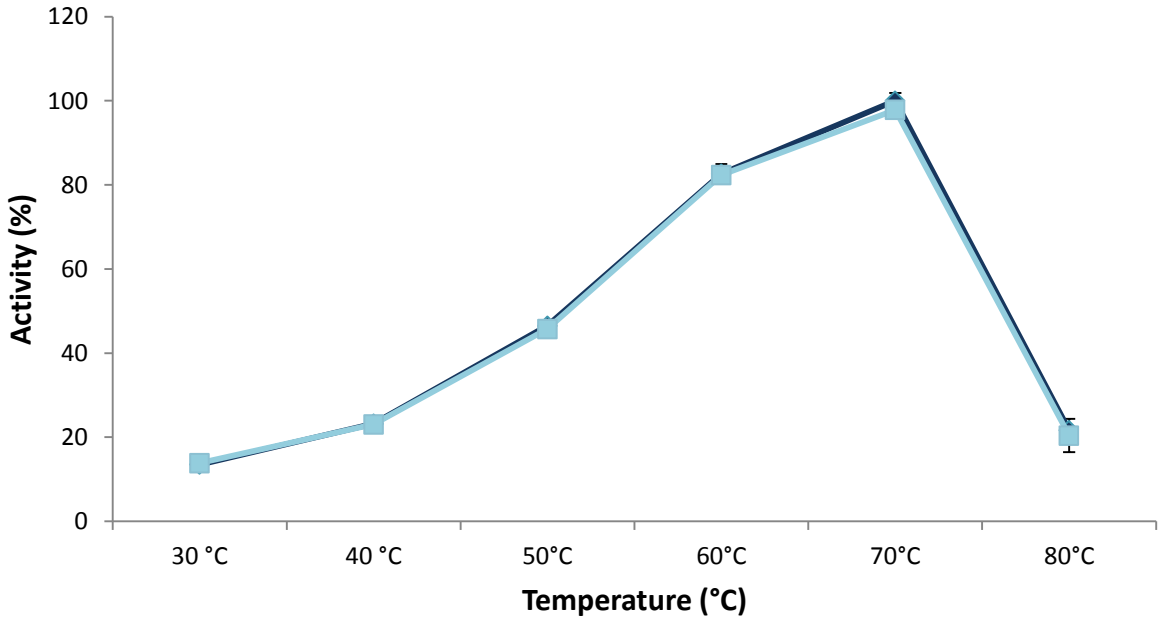


Figure 2.9: Temperature dependency of *Gsp*DAH7PS^{WT} showing CM activity at different temperatures (30 °C, 40 °C, 50 °C, 60 °C, 70 °C and 80 °C) without the addition of any metal ions. Substrate (70 µM CA) buffer (50 mM BTP, pH 6.8) and enzyme (3 µL at 13 mg/mL) concentrations were kept constant. Enzyme was incubated at desired temperature with the reaction mixture for either one minute (shown in dark blue) or five minutes (shown in light blue).

The results show that *Gsp*DAH7PS^{WT} CM enzyme activity increases as temperature increases. The wild-type enzyme is most active at 70 °C, and then activity drops sharply at

80 °C. This is consistent with the wild-type DAH7PS enzymatic activity that also drops at 80 °C (section 2.6). Incubating the wild-type enzyme with the reaction mixture for either one or five minutes did not make an appreciable difference to the data (Figure 2.9).

2.8 Michaelis-Menten kinetics

2.8.1 DAH7PS activity

Once the most activating metal ion (section 2.5) and temperature (section 2.6) were determined for *Gsp*DAH7PS^{WT}, the results were incorporated in the experimental procedure to obtain the steady state kinetic parameters of the enzyme. The steady state kinetic parameters were measured by monitoring the consumption of PEP at 232 nm (Figure 2.10). The continuous spectrophotometric assay for the enzyme was carried out at 60 °C, despite this being lower than the temperature *Gsp*DAH7PS^{WT} is most active at (section 2.6), this ensured sufficient time for accurate rates to be recorded *in vitro* without the degradation of E4P, used to initiate the continuous assay, due to its instability at elevated temperatures.[1] The use of a continuous assay (previously used in the characterisation of *Tma*DAH7PS[68] and *Pfu*DAH7PS[40]) instead of a discontinuous assay, is a more convenient method that gives more accurate data, as more data points can be obtained.[40] Other studies characterising type Iβ DAH7PS enzymes such as *Bsu*DAH7PS,[13] *Pgi*DAH7PS[12] and *Ape*DAH7PS[15] used a discontinuous colourimetric or Aminoff periodate-thiobarbituric acid assay system, or a combination of both continuous and discontinuous methods.

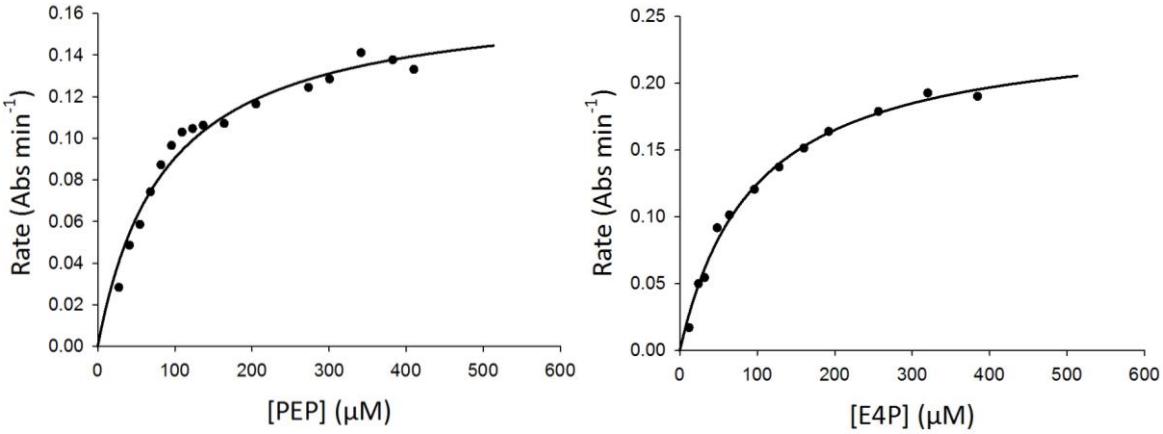


Figure 2.10: Michaelis-Menten plots obtained for *Gsp*DAH7PS^{WT} showing the apparent PEP (left) K_m value of 87 ± 9 µM and E4P (right) K_m with a value of 95 ± 8 µM. A k_{cat} value of 45 ± 4 s^{-1} was obtained. Standard assays were carried out at 60 °C in 50 mM BTP buffer (pH 6.8) and 100 µM Cd^{2+}.

Initial velocities for the catalytic activity of *Gsp*DAH7PS^{WT} were determined by keeping the concentration of one substrate fixed (PEP concentration was fixed at 338 µM and E4P concentration was fixed at 329 µM) while varying the concentration of the other substrate. The apparent K_m values for PEP and E4P were found to be 87 ± 9 µM and 95 ± 8 µM respectively, with a calculated k_{cat} value of 45 ± 4 s^{-1} (

Table 2.1). Due to the difference in temperatures and assay conditions used to obtain kinetic parameters of other type Iβ DAH7PS enzymes, it is difficult to make accurate comparisons of the data obtained. It is noted however that *Gsp*DAH7PS^{WT} does have a larger k_{cat} value in comparison to the other type Iβ DAH7PS enzymes.

Table 2.1: Kinetic parameters for *Gsp*DAH7PS^{WT} and other type Iβ DAH7PS enzymes

Enzyme	K_m^{PEP} (µM)	K_m^{E4P} (µM)	k_{cat} (s^{-1})	k_{cat}/K_m^{PEP} (µM.s^{-1})	k_{cat}/K_m^{E4P} (µM.s^{-1})
***Gsp*DAH7PS^{WT}**	87 ± 9	95 ± 8	45 ± 4	0.52	0.47
***Bsu*DAH7PS[13]**	139 ± 11.4	1760 ± 110	4.6 ± 0.1	3.3 x 10^{-2}	2.6 x 10^{-3}
***Pgi*DAH7PS[12]**	421 ± 43	1238 ± 141	1.6 ± 0.1	3.8 x 10^{-3}	1.3 x 10^{-3}
***Tma*DAH7PS[68]**	4.9 ± 0.4	13 ± 1	11.7 ± 0.2	2.4	0.9
***Pfu*DAH7PS[40]**	120 ± 20	28 ± 4	1.5 ± 0.1	1.3 x 10^{-2}	5.4 x 10^{-2}
***Ape*DAH7PS[15]**	890 ± 30	280 ± 20	1.0 ± 0.2	1.1 x 10^{-3}	3.6 x 10^{-3}

2.8.2 Chorismate mutase activity

The steady state kinetic parameters for the CM activity of *Gsp*DAH7PS^{WT} were measured by tracking the consumption of CA at 274 nm (Figure 2.11). While the temperature dependency study for *Gsp*DAH7PS^{WT} showed that the enzyme's enzymatic activity was highest at 70 °C (Section 2.7), assays were carried out at 50 °C. This was to ensure that CA would not degrade while carrying out the experiment, as it is a highly unstable compound, especially at elevated temperatures.

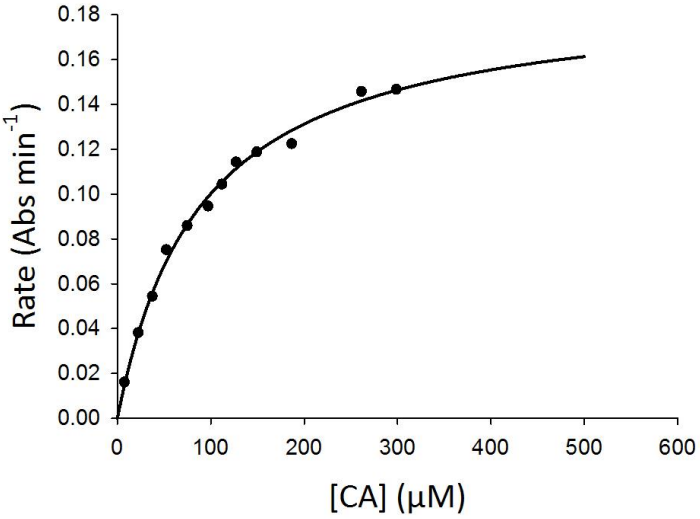


Figure 2.11: Michaelis-Menten plot obtained for *Gsp*DAH7PS^{WT} showing the apparent CA K_m with a value of 88 ± 6 µM and a k_{cat} value of 1.9 ± 0.13 s^{-1}. Standard assays were carried out at 50 °C in 50 mM BTP buffer (pH 6.8) without the addition of any metal ions.

Initial velocities for the catalytic CM activity were measured by initiating the equilibrated reaction mixture with varying concentrations of CA. The apparent K_m value for CA was 88 ± 6 µM for with a calculated k_{cat} value of 1.9 ± 0.13 s^{-1} (Table 2.2). In comparison to *Bsu*DAH7PS, *Gsp*DAH7PS^{WT} displays a much more active CM domain with a higher affinity for CA (Table 2.2).

Table 2.2: Kinetic parameters for CM activity of *Gsp*DAH7PS^{WT} and other DAH7PS enzymes

Enzyme	K_m^{CA} (μM)	k_{cat} (s^{-1})	k_{cat}/K_m^{CA} (μM.s^{-1})
***Gsp*DAH7PS^{WT}**	88 ± 6	1.9 ± 0.13	$2.1x10^{-2}$
***Bsu*DAH7PS[13]**	851 ± 97	0.41 ± 0.01	$4.8x10^{-4}$
***Pgi*DAH7PS[12]**	Not active	…….	…….
***Mtu*CM+*Mtu*DAH7PS[63]**	34 ±3	8.1 ± 1.9	0.24
***Eco*DAH7PS[69]**	30 ± 1	41.4 ± 0.4	1.4

2.9 Inhibition

Characterisation studies have shown that some DAH7PS enzymes have evolved extensions to the core $(\beta/\alpha)_8$ barrel that are involved in feedback regulation. [16,46,48,49] Previously characterised DAH7PS enzymes fused to an allosteric CM domain demonstrated inhibition of the DAH7PS activity in the presence of CA or PA.[12] These results led to the hypothesis that *Gsp*DAH7PS^{WT} may also be inhibited in the presence of the branch point intermediates, CA or PA, due to the presence of an N-terminal CM domain extension.

An inhibition study was carried out on *Gsp*DAH7PS^{WT} in the presence of PA to determine if this compound would inhibit the enzyme. DAH7PS activity was tested using the same Method as above (section 2.8.1). Substrate (PEP, E4P), metal (Cd^{2+}), and enzyme concentrations were kept constant in 50 mM BTP buffer (pH 6.8), while the concentration of PA was varied. PA, rather than CA, was the inhibitor chosen since CA is unstable at elevated temperatures. CA is also converted to PA via CM almost instantaneously, which would be occurring while the experiment is being carried out, leading to lack of clarity about the inhibitory species in the assay. Furthermore, it has been previously suggested in a study on *Bsu*DAH7PS, that PA is the true inhibitor of the DAH7PS domain instead of CM in *vivo.*[12] This was due to a sigmoid inhibition curve being obtained for *Bsu*DAH7PS in the presence of increasing concentrations of CA. It is believed that CA was being converted to PA initially, which in turn inhibited *Bsu*DAH7PS activity.

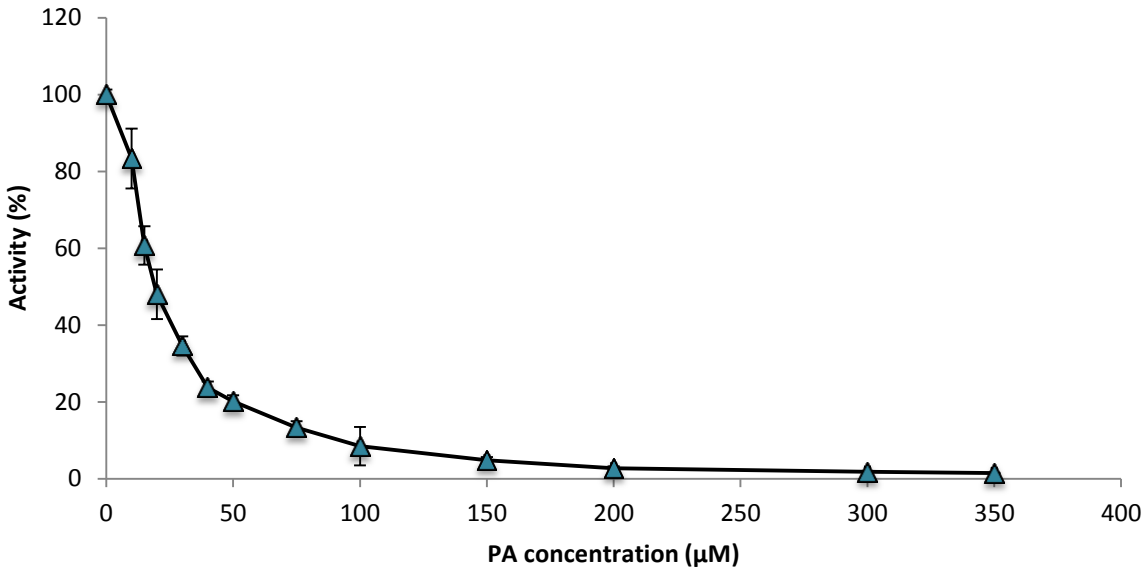


Figure 2.12: Response of *Gsp*DAH7PS^{WT} activity to increasing concentrations of PA. Standard assays were carried out at 60 °C, with 121 µM PEP and 123 µM E4P in 50 mM BTP buffer (pH 6.8). The data points were measured in triplicate with error bars indicating the standard deviation.

The results clearly demonstrate that PA inhibits *Gsp*DAH7PS^{WT} (Figure 2.12). A logistic function fitted to the data gives an IC50 value of 25.3 µM. This result shows increased sensitivity to PA in comparison to *Bsu*DAH7PS (IC50 value of 100 µM), but significantly less sensitivity to PA in comparison to *Pgi*DAH7PS, where the CM domain is fused to the C-terminal of the DAH7PS barrel (IC50 value of 1.5 µM) (Table 2.3).

Table 2.3: Estimated IC50 values of *Gsp*DAH7PS^{WT}, *Bsu*DAH7PS and *Pgi*DAH7PS

Enzyme	**Estimated IC50 (µM)**	
	Prephenate	Chorismate
***Gsp*DAH7PS^{WT}**	25.3	not determined
***Bsu*DAH7PS[12]**	100	1000
***Pgi*DAH7PS[12]**	1.5	2.5

2.10 Analytical ultracentrifugation of *Gsp*DAH7PS^{WT}

Gel filtration traces of *Gsp*DAH7PS^{WT} showed that the enzyme eluted as a single peak containing a shoulder. SDS-PAGE analysis along with activity assays determined that both the peak and the shoulder contained *Gsp*DAH7PS^{WT}. This suggested that the wild-type enzyme might be in equilibrium between two different oligomeric states in solution.

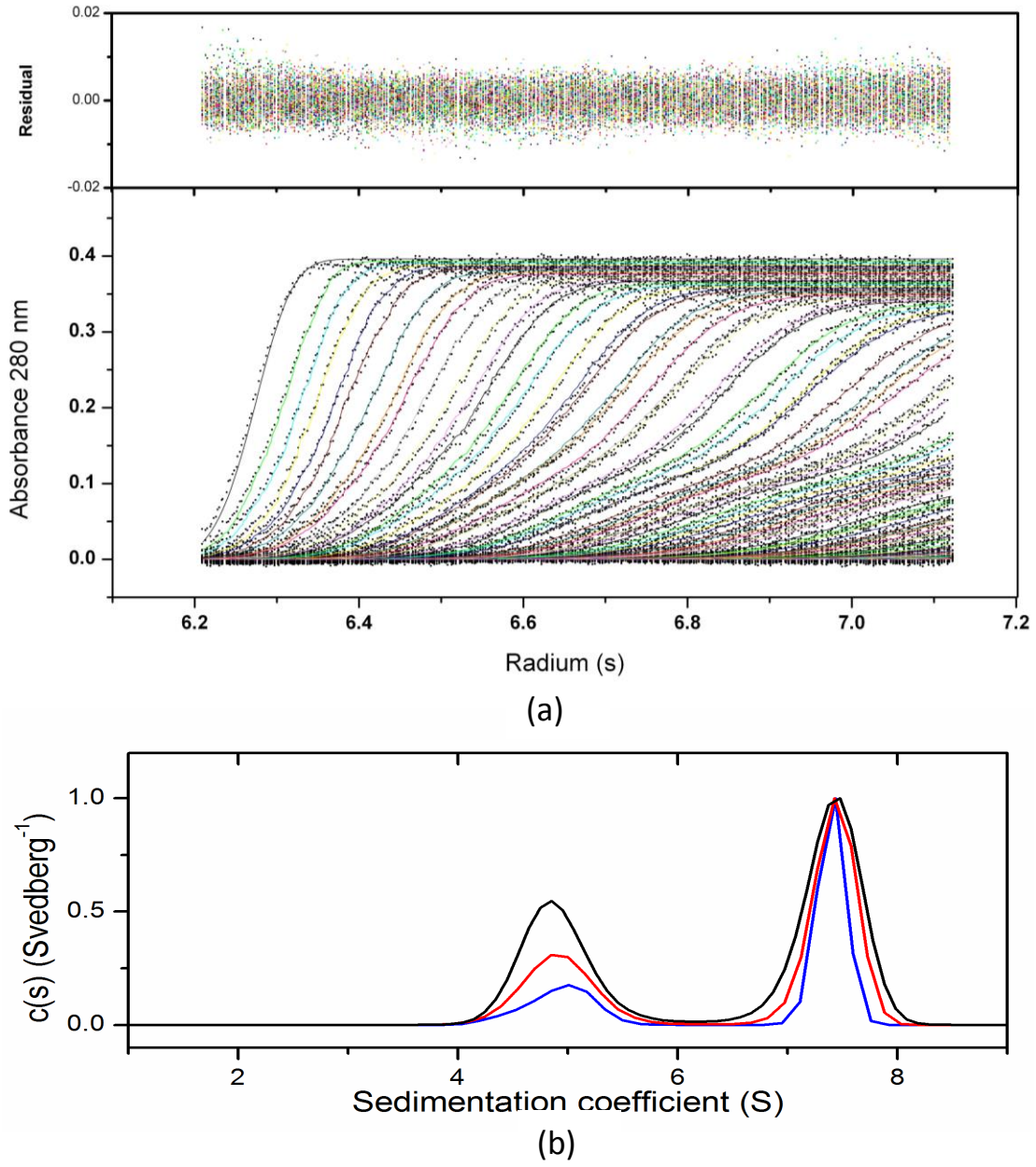


Figure 2.13: Sedimentation velocity analysis of *Gsp*DAH7PS^{WT} at 1.8 mg/ml (blue line), 1 mg/ml (red line) and 0.5 mg/ml (black line); all curves were normalised to 1. Experiments were carried in SEC buffer (pH 7.4 at 20 °C). Data was collected at 50,000 rpm at a wavelength of 280 nm. (A) the best fit and residuals of the experimental data. (B) the distribution of the species.

An investigation into the quaternary assembly of *Gsp*DAH7PS^{WT} in solution was carried out using analytical ultracentrifugation (AUC) (Figure 2.13). By applying extreme gravitational forces on an enzyme of interest, and tracking its sedimentation behaviour, this technique helps to determine if the enzyme exists as one or multiple species in solution. The technique also helps to determine whether quaternary structure of the enzyme is concentration dependent.

Stock solutions of wild-type enzyme were buffer exchanged into freshly made SEC buffer (10 mM BTP, pH 7.4, 40 mM KCl, 200 µM PEP and 10 µM EDTA) and made up to three different concentrations (1.8 mg/mL, 1 mg/mL and 0.5 mg/mL). In an ultracentrifuge, each sample was run simultaneously with a blank containing an exact buffer match to ensure a precise profile of the enzyme in solution was obtained.

According to the absorbance versus radial position data, two sedimentation boundaries exist for *Gsp*DAH7PS^{WT}, suggesting that the enzyme exists as two oligomeric species in solution. A best fit of the experimental data showed the presence of two species that had sedimentation coefficients of 4.9 S and 7.3 S. Fitting the sedimentation velocity data, correlated to calculated masses of 80 kDa and 145 kDa for *Gsp*DAH7PS^{WT}. This is consistent with the molecular masses for the dimeric (83 kDa) and tetrameric (165 kDa) assemblies of *Gsp*DAH7PS^{WT}.

To ensure that equilibrium between two oligomeric states was being observed, sedimentation equilibrium experiments at different concentrations were carried out. Normalising the tetrameric sedimentation coefficient boundaries obtained to 1 for all concentrations showed that as the concentration of enzyme increased, tetramer formation was favoured with a reduced dimer formation.

CA is converted to PA in the shikimate pathway, a reaction mediated by CM. PA has been shown to completely inhibit the DAH7PS domain of the wild-type enzyme, and this is likely to arise from it binding to the CM domain (section 2.9).

The same experiment as above was carried out with the wild-type enzyme in the presence of PA, in attempt to find any effects of the inhibitor on the dimer tetramer equilibrium. Such information could give key information on the mechanism for inhibition of *Gsp*DAH7PS^{WT} by PA. In this experiment, a sample was run at a concentration of 1 mg/ml in the presence of 500 µM PA (Figure 2.14). Data obtained was compared with the AUC data for *Gsp*DAH7PS^{WT} in the absence of PA at a concentration of 1 mg/ml.

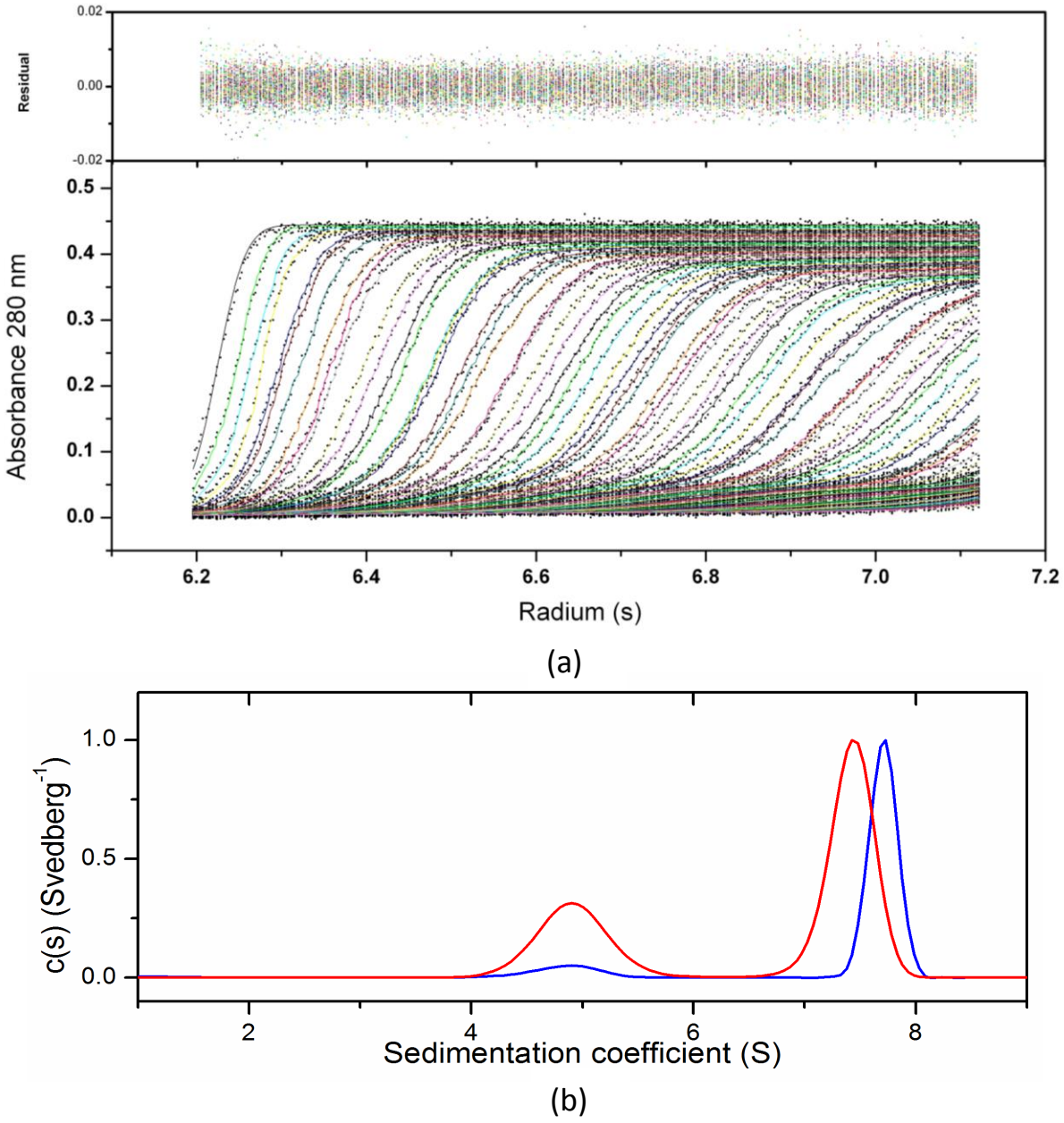


Figure 2.14: Sedimentation velocity analysis of *Gsp*DAH7PS^{WT} at a concentration of 1 mg/ml in the presence of 500 µM PA (blue line), and in the absence of PA (red line). Experiments were carried in SEC buffer (pH 7.4 at 20 °C). Data was collected at 50,000 rpm at a wavelength of 280 nm. (a) the best fit of the experimental data. (b) the distribution of the species.

Absorbance versus radial position data showed the existence of two sedimentation boundaries for the wild-type enzyme in the presence of PA, suggesting that the enzyme still exists as two oligomeric species in solution. A best fit of the experimental data showed the presence of two species that had sedimentation coefficients of 4.5 S and 7.8 S. Fitting the sedimentation velocity data, correlated to calculated masses of 80 kDa and 170 kDa for *Gsp*DAH7PS^{WT}. This is consistent with the molecular masses for the dimer (83 kDa) and tetramer (165 kDa) of *Gsp*DAH7PS^{WT}.

Comparison of the sedimentation velocity data in the presence and absence of PA by normalising the tetrameric sedimentation boundaries to 1, showed that addition of PA lead to a dramatic increase in tetramer formation and a reduction in dimer. PA binding to the regulatory CM domain leads to the formation of a more stable quaternary enzyme where CM domains from diagonally opposite sides of the tetramer interlock with one another to hold the tetrameric structure together.

2.11 Summary

*Gsp*DAH7PS^{WT} was expressed and purified successfully by heat treatment, metal-affinity and size exclusion chromatography. Mass spectrometry determined that the enzyme was of the expected mass. The enzyme was found to be thermally stable and demonstrated both CM and DAH7PS activity.

DSF results showed that different metal ions had different stabilising effects on *Gsp*DAH7PS^{WT} in comparison to EDTA, which was destabilising to the enzyme. The substrates PA and CA were also destabilising to the enzyme.

*Gsp*DAH7PS^{WT} was determined to be affected by temperature and to be metal dependent, with the highest CM and DAH7PS enzymatic activity's demonstrated at 70 °C, and with Cd^{2+} being the most activating metal ion for the enzyme.

Kinetic assays revealed the Michaelis-Menten kinetics of the wild-type enzyme for both the CM and DAH7PS domain activity's. The DAH7PS domain had apparent K_m values of 87 ± 9 µM and 95 ± 8 µM for PEP and E4P respectively, with a calculated k_{cat} value of 45 ± 4 s^{-1}, a value higher than has been reported for other characterised type Iβ DAH7PS enzymes. The CM domain had an apparent K_m value of 88 ± 6 µM for CA with a calculated k_{cat} value of 1.9 ± 0.1 s^{-1}.

AUC determined that *Gsp*DAH7PS^{WT} exists in equilibrium between a dimeric and tetrameric state with the equilibrium shifting toward tetramer formation with increasing enzyme concentration. PA, which was found to completely inhibit DAH7PS activity, led to a dramatic increase in tetramer formation and a decrease in dimer formation.

Based on previous studies of type Iβ and other regulated DAH7PS enzymes[12,16] it is suggested that the N-terminal CM of *Gsp*DAH7PS^{WT} is for allosteric regulation. To determine if this is really the case, further work was needed. The next step was to carry out a domain truncation on *Gsp*DAH7PS^{WT} by separating the catalytic and putative regulatory domains in order to characterise them independently. This would help determine if the CM domain is essential for enzyme regulation.

Chapter 3

Expression and biochemical characterisation of truncated *Geobacillus sp* CM and DAH7PS

3.1 Overview

Previous studies of DAH7PS enzymes over the years have shown that all DAH7PS enzymes from different organisms share the same catalytic core $(\beta/\alpha)_8$ barrel fold.[16] Some of these enzymes were found to be regulated, containing different small domain extensions fused to the core barrel, and these proteins are believed to have evolved from an undecorated and unregulated ancestral DAH7PS enzyme.[20] Amongst the type Iβ DAH7PS subfamily are enzymes that consist of only the $(\beta/\alpha)_8$ barrel such as *Pfu*DAH7PS[40] and *Ape*DAH7PS,[15] both of which have no known regulators or natural inhibitors. These enzymes differ from others in the type Iβ DAH7PS subfamily such as *Bsu*DAH7PS,[13] *Pgi*DAH7PS[12] and *Tma*DAH7PS,[68] which contain extensions to the catalytic barrel and are regulated. It has been suggested that these extensions were incorporated for the sole purpose of feedback regulation, where the activity of the catalytic DAH7PS domain is regulated through pathway intermediates or end products binding to the extensions. Type Iβ DAH7PS enzymes are the most widely distributed in nature and are therefore believed to be most closely related to the common ancestral DAH7PS enzyme, making them a great deal of importance in determining the evolution pattern of the DAH7PS family.[19]

*Gsp*DAH7PS^{WT} is a type Iβ DAH7PS enzyme containing an N-terminal CM domain, believed to have evolved for the sole purpose of regulation of catalytic barrel, and while it has shown that this enzyme is regulated by the pathway intermediate PA, further investigation is needed to determine this. In this chapter, a domain truncation was carried out on the wild-

type enzyme leading to independent expression and purification of the DAH7PS and CM domains. Each domain was characterised in a similar manner to *Gsp*DAH7PS^{WT}, and comparative studies were carried out between the wild-type enzyme and the truncated variants to help determine the effects of this.

3.2 Truncation site

*Gsp*DAH7PS^{WT} was truncated between residues Lys92 and Ala93 located at the linker region connecting the DAH7PS domain and the CM domain. Previous domain truncation studies,[12,16] structural and sequence comparisons with *Pfu*DAH7PS and *Lmo*DAH7PS (Figure 3.1) helped to determine the linker region, and a truncation at this position appeared to be the best choice as it is most likely to lead to separate and intact DAH7PS domain. It should be noted that although the initial analysis was carried out prior to the availability of the crystal structure of this protein, the structure confirmed the linker position (Figure 3.2).

```
GspDAH7PS       MGNERLDELRARVDEINLQLLKLINERGRLVQEIGKIKEAQGTHRYDPVRERKMLDLISE 60
LmoDAH7PS       MVNTNLEELRTQVDQLNIDLLELISKRANLVQEIGKIKGTQGSLRFDPLREREMLNTILA 60
PfuDAH7PS       ------------------------------------------------------------

GspDAH7PS       HNDGPFETSTLQHIFKEIFKAALELQEDDHRKALLVSRKKHPENTIVEVKGERIGDGNQY 120
LmoDAH7PS       ANEGPFEDSTVQKLFKEIFKAGLELQEEDHSKALLVSRKNKKEDTIVTVKGLPIGNGEPV 120
PfuDAH7PS       ----------------------------------MKYSKEYKEKTVVKINDVKFGEG-FT 25
                                                    :   *:  *.*:* ::.  :*:*

GspDAH7PS       FVMGPCAVESYEQVAAVAEAVKKQGIKLLRGGAYKPRTSPYDFQGLGVEGLKILKRIADE 180
LmoDAH7PS       FVFGPCSVESYEQVAAVAESIKAKGLKLIRGGAFKPRTSPYDFQGLGLEGLKILKRVSDE 180
PfuDAH7PS       IIAGPCSIESRDQIMKVAEFLAEVGIKVLRGGAFKPRTSPYSFQGYGEKALRWMREAADE 85
                :: ***::** :*:  *** :   *:*::****:*******.*** * :.*: ::. :**

GspDAH7PS       FDLAVISEIVTPADIEIALDYIDVIQIGARNMQNFELLKAAGQVNKPILLKRGLAATIEE 240
LmoDAH7PS       YGLGVISEIVTPADIEVALDYVDVIQIGARNMQNFELLKAAGRVDKPILLKRGLSATIEE 240
PfuDAH7PS       YGLVTVTEVMDTRHVELVAKYSDILQIGARNSQNFELLKEVGKVENPVLLKRGMGNTIQE 145
                :.* .::*:: . .:*:. .* *::****** ******* .*:*::*:*****:. **:*

GspDAH7PS       FINAAEYIMSQGNGQIILCERGIRTYERATRNTLDISAVPILKKETHLPVFVDVTHSTGR 300
LmoDAH7PS       FIGAAEYIMSQGNGKIILCERGIRTYEKATRNTLDISAVPILKKETHLPVMVDVTHSTGR 300
PfuDAH7PS       LLYSAEYIMAQGNENVILCERGIRTFETATRFTLDISAVPVVKELSHLPIIVDPSHPAGR 205
                :: :*****:*** ::*********:* *** ********::*: :***::** :*.:**

GspDAH7PS       RDLLIPCAKAALAIGADGVMAEVHPDPAVALSDSAQQMDIAQFNEFMEEVRAFQRQFVRA 360
LmoDAH7PS       KDLLLPCAKAALAIEADGVMAEVHPDPAVALSDSAQQMDIPEFEEFWNAILASN-LVPHK 359
PfuDAH7PS       RSLVIPLAKAAYAIGADGIMVEVHPEPEKALSDSQQQLTFDDFLQLLKELEALGWKG--- 262
                :.*::* **** ** ***:*.****:*  ***** **: : :* :: : : *

GspDAH7PS       LE 362
LmoDAH7PS       IK 361
PfuDAH7PS       --

                              * Conserved residues
                              : Conserved substitutes
                              . Semi-conservative substitutes
```

Figure 3.1: ClustalW sequence alignment of *Gsp*DAH7PS^{WT}, *Lmo*DAH7PS and *Pfu*DAH7PS which consists of only a catalytic DAH7PS domain. The linker residues connecting the N-terminal CM domain and the DAH7PS domain of *Lmo*DAH7PS are highlighted in yellow. The truncation site of *Gsp*DAH7PS^{WT} is highlighted in yellow. The CM and DAH7PS truncated mutant sequences of *Gsp*DAH7PS^{WT} are shown in green and red respectively.

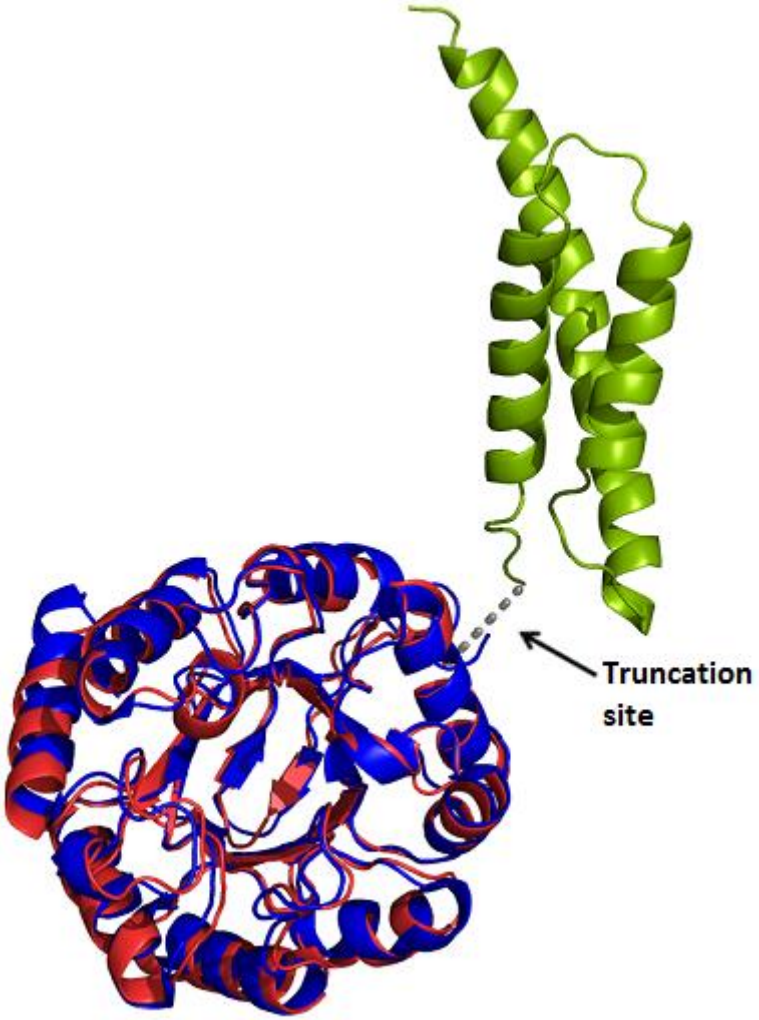


Figure 3.2: Structural comparison of *Gsp*DAH7PS^{WT} and *Pfu*DAH7PS (PDB code 1ZCO) subunits showing the truncation site of *Gsp*DAH7PS^{WT}. The *Gsp*DAH7PS^{WT} monomer is shown in red to represent the catalytic DAH7PS domain and green to represent the regulatory N-terminal CM domain. The grey dotted line represents the missing linker region (not observable due to high flexibility of the linker) connecting both domains of the *Gsp*DAH7PS^{WT} monomer. The *Pfu*DAH7PS monomer is shown in dark blue to represent the DAH7PS domain.

The genes corresponding to the ORFs for both *Gsp*DAH7PS^{DAH7PS} (not including the sequence encoding the non-cleavable C-terminal tag) and *Gsp*DAH7PS^{CM} were amplified from the plasmid, which contained the gene for *Gsp*DAH7PS^{WT} using primers and a two-round nested PCR protocol (experimental Section 5.3.1) (Figure 3.3).

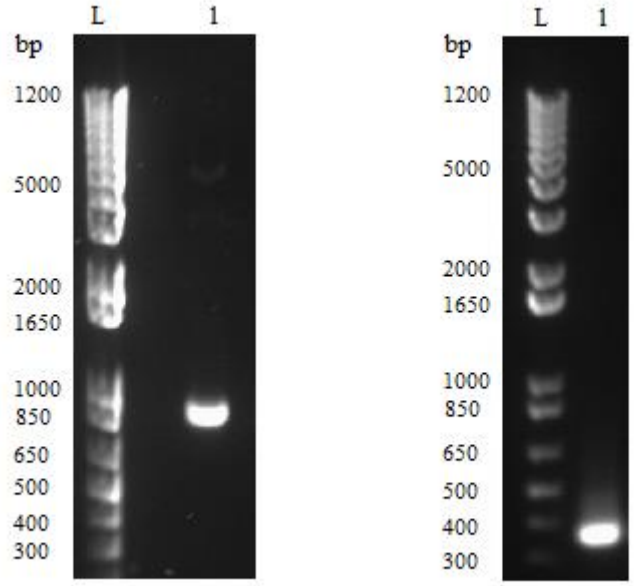


Figure 3.3: Agarose gels showing round 2 nested PCR gene products coding for (a) *Gsp*DAH7PS^{DAH7PS} and (b) *Gsp*DAH7PS^{CM} in lane 1 of each gel.

3.3 Cloning and expression of *Gsp*DAH7PS^{DAH7PS}

Once amplified, the gene encoding for *Gsp*DAH7PS^{DAH7PS} was ligated into the entry vector pDONR221 using a recombination reaction catalysed by BP clonase. DNA sequencing confirmed that the gene of interest with the correct sequence had been inserted. The gene coding for *Gsp*DAH7PS^{DAH7PS} was then transferred to the destination vector pDEST17 using a further recombination reaction catalysed by LR clonase. This vector allows the desired protein to be expressed with a hexaHis-tag at the N-terminal, making it easy to detect and purify. The tag includes a Tobacco Etch Virus (TEV) protease cleavage site that allows for the removal of the tag using TEV protease following protein purification.

The pDEST17 vector containing the gene for His-*Gsp*DAH7PS^{DAH7PS} was transformed into *E. coli* chaperone 3 cells. Cells were grown and protein expression was induced by the addition of IPTG. The cells were harvested from the growth media by centrifugation, resuspended in HisTrap HP binding buffer (20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole, pH 7.4), lysed, heat treated at 55˚C for 30 minutes and left to cool to room temperature. The

soluble protein of interest was separated from the insoluble material and cell debris by centrifugation. SDS-PAGE analysis of the cells before and after induction, and the supernatant after lysis and centrifugation showed that His-*Gsp*DAH7PS^{DAH7PS} was overexpressed at the predicted molecular weight (approximately 31 kDa) with most of the protein being present in the supernatant (Figure 3.4). A basic activity assay confirmed that the supernatant had DAH7PS activity.

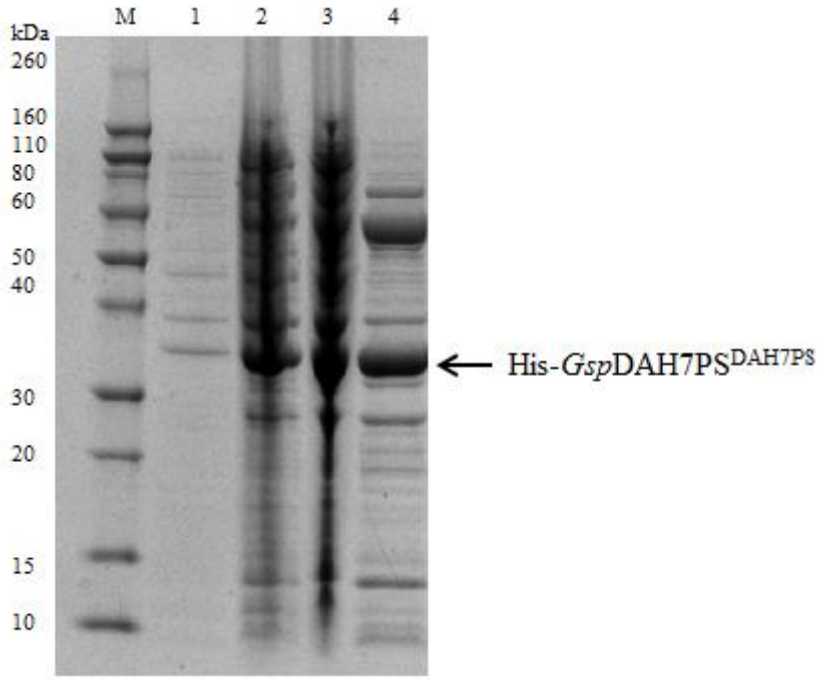


Figure 3.4: SDS-PAGE analysis for His-*Gsp*DAH7PS^{DAH7PS} before induction (lane 1), after induction (lane 2), after lysis and heat treatment (lane 3), and in the supernatant after centrifugation (lane 4).

3.4 Purification of *Gsp*DAH7PS^{DAH7PS} variant

His-*Gsp*DAH7PS^{DAH7PS} was further purified using nickel ion affinity chromatography. Following binding of the protein in low concentration imidazole buffer (20 mM), the protein was eluted with a high concentration of imidazole using a linear gradient elution. SDS-PAGE analysis helped to determine the eluted fractions containing protein of interest, which were pooled, concentrated and put down a desalting column to reduce the high imidazole concentration (Figure 3.5). TEV protease was added to cleave the His-tag from the protein,

and left overnight at 4 °C. The protein and His-tag mixture were then passed through a nickel ion affinity column another time to separate the His-tag from the untagged *Gsp*DAH7PS^{DAH7PS}, which was eluted in the unbound fractions (Figure 3.5). Fractions containing *Gsp*DAH7PS^{DAH7PS} were pooled, concentrated and further purified using size exclusion chromatography (SEC). *Gsp*DAH7PS^{DAH7PS} was found to elute in two peaks from the SEC column (Figure 3.6).

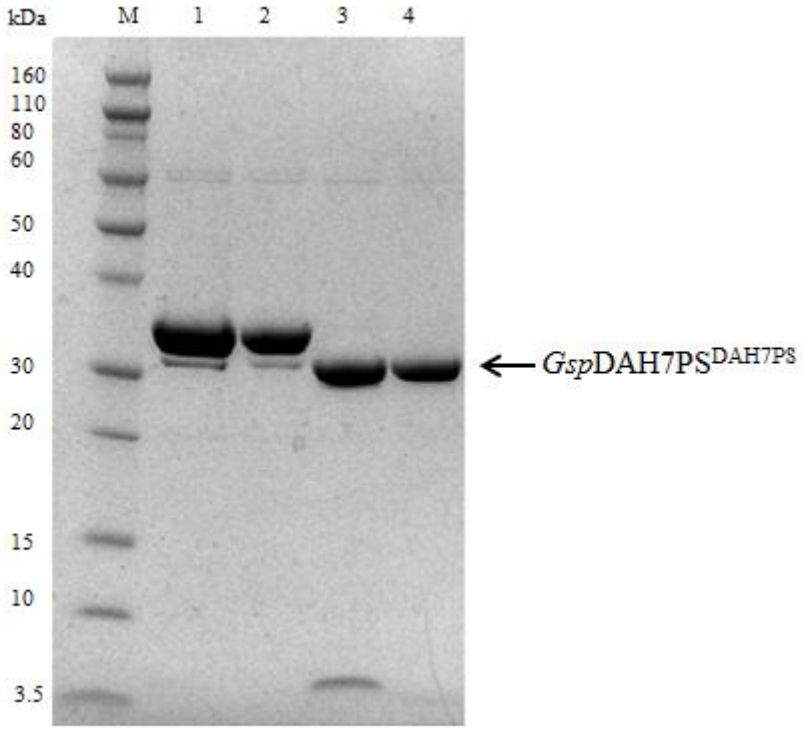


Figure 3.5: SDS-PAGE analysis for *Gsp*DAH7PS^{DAH7PS} after first nickel ion affinity column (lane 1), after desalting column (lane 2), after TEV cleavage (lane 3) and after second nickel ion affinity column (lane 4).

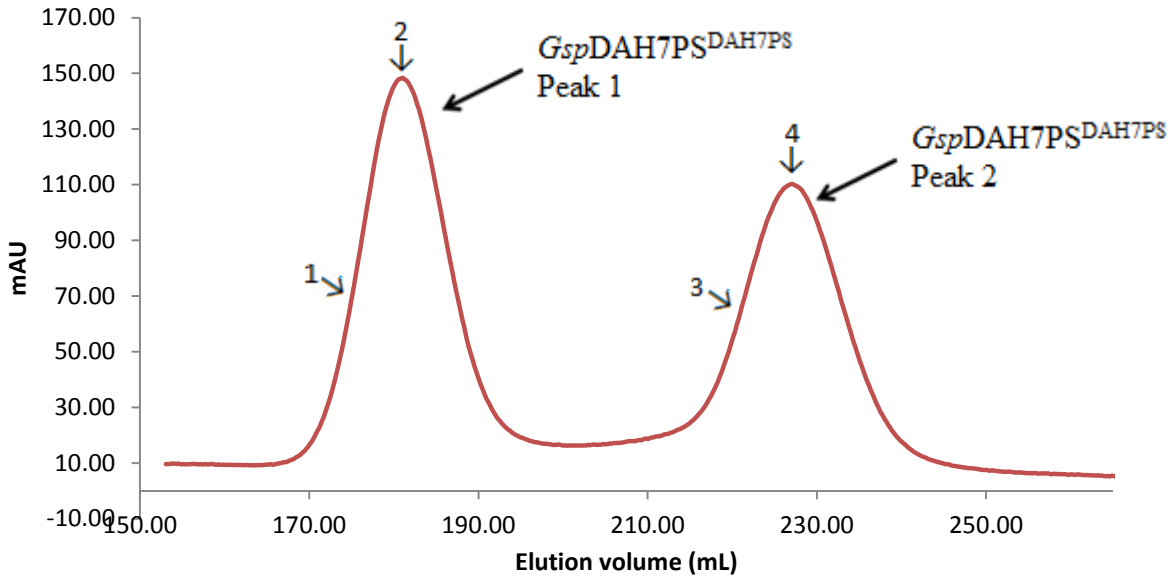


Figure 3.6: SEC trace of eluted *Gsp*DAH7PS^{DAH7PS} in SEC buffer (10 mM BTP, pH 7.4, 40 mM KCl, 200 µM PEP and 10 µM EDTA) showing both peaks corresponding to the eluted protein. The numbers 1 to 4 represent the lane number on the SDS-PAGE gel of where the fractions from the SEC run were taken (Figure 3.7).

SDS-PAGE analysis showed that both peaks contained protein of the right molecular mass, with protein from each peak containing DAH7PS activity according to an enzyme activity assay (Figure 3.7). Fractions from each peak were pooled separately, concentrated, put into aliquots, and stored at -80 °C

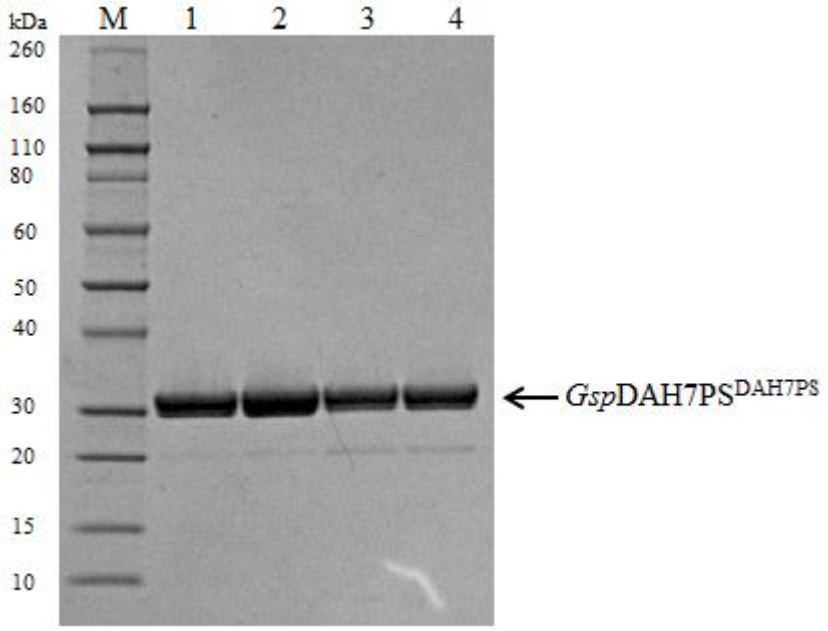


Figure 3.7: SDS-PAGE gel of final collected fractions after SEC of *Gsp*DAH7PS^{DAH7PS}. Lanes 1 and 2 showing fractions collected from the first peak. Lanes 3 and 4 showing fractions collected from the second peak (Figure 3.6)

The presence of two peaks on the SEC trace along with enzyme from each peak being active and of the correct molecular weight, suggested that *Gsp*DAH7PS^{DAH7PS} is present in two oligomeric states in solution. Further experimental tests were needed to confirm this.

3.5 Cloning and expression of *Gsp*DAH7PS^{CM}

The gene coding for *Gsp*DAH7PS^{CM} was ligated into the entry vector pDONR221 using BP clonase. DNA sequencing was carried out to confirm that the gene of interest with the correct sequence had been inserted. The plasmid bearing the *Gsp*DAH7PS^{CM} gene was initially recombined with the destination vector pDEST 17 using an LR clonase catalysed reaction, to add an N-terminal hexaHis-tag. Attempts were made to purify the enzyme, and while there was expression, the enzyme showed no signs of solubility. A different destination vector had to be used to aid in the solubility of *Gsp*DAH7PS^{CM}.

The plasmid bearing the *Gsp*DAH7PS^{CM} gene was then recombined with the destination vector pDEST15, allowing the desired protein to be expressed with an N-terminal GST-tag

that is cleavable with TEV protease. The GST protein aids with solubility of the desired protein and helps deliver higher protein yields. This plasmid was transformed into *E. coli* One Shot BL21 star (DE3) cells (Life Technologies) for expression.

Cells bearing the plasmid encoding the GST-*Gsp*DAH7PS^{CM} gene were grown and protein expression was induced. The cells were harvested, resuspended in GSTrap binding buffer PBS, (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 pH 7.3) and then lysed before being centrifuged in order to separate the soluble protein of interest from the remaining lysed cells. SDS-PAGE analysis of the whole cells before and after induction, along with the supernatant showed that GST-GspDAH7PS^{CM} was overexpressed at the predicted molecular weight (approximately 36 kDa) with most of the protein being present in the supernatant (Figure 3.8). An activity assay indicated that there was CM activity in the supernatant.

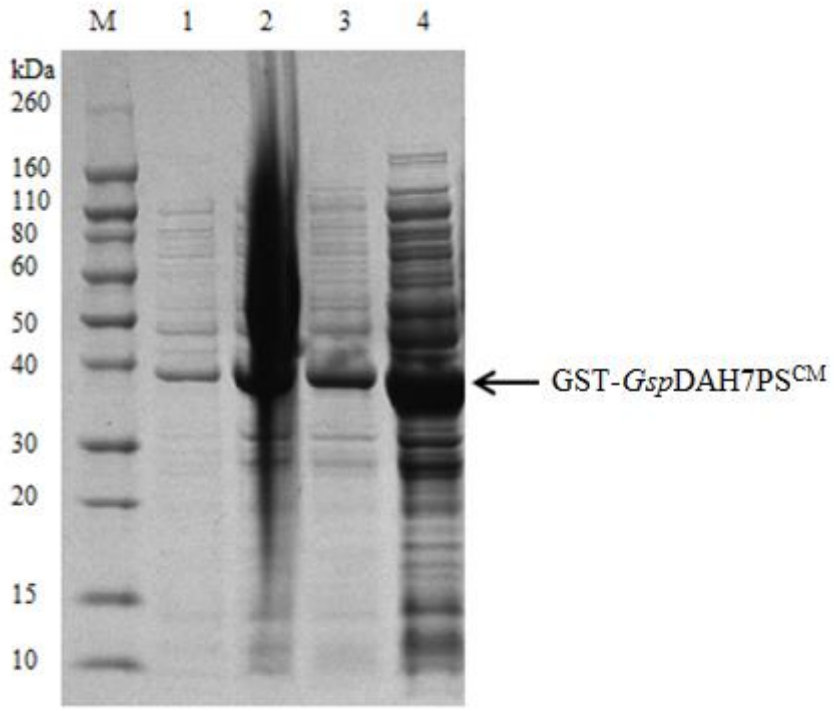


Figure 3.8: SDS-PAGE analysis for GST-*Gsp*DAH7PS^{CM} before induction (lane 1), after induction (lane 2), whole cell before centrifugation (lane 3), and in the supernatant after centrifugation (lane 4).

3.6 Purification of *Gsp*DAH7PS^{CM} variant

The collected supernatant was applied to a GSTrap affinity column. Column bound GST-*Gsp*DAH7PS^{CM}, was then eluted with a high concentration of glutathione using a linear gradient. Determined by SDS-PAGE analysis, eluted fractions containing GST-*Gsp*DAH7PS^{CM} were then were then collected, concentrated and passed through a desalting column to remove the high glutathione content. TEV protease was added to cleave and separate the GST-tag from the protein (Figure 3.9). The resulting supernatant was put down the GSTrap again to separate the GST-tag, which binds to the GSTrap column, from *Gsp*DAH7PS^{CM} that was eluted in the non-bound fractions. *Gsp*DAH7PS^{CM} fractions were pooled, concentrated and further purified using SEC. *Gsp*DAH7PS^{CM} was found to elute in three peaks off the SEC column along with a shoulder to the left of the second peak. (Figure 3.10)

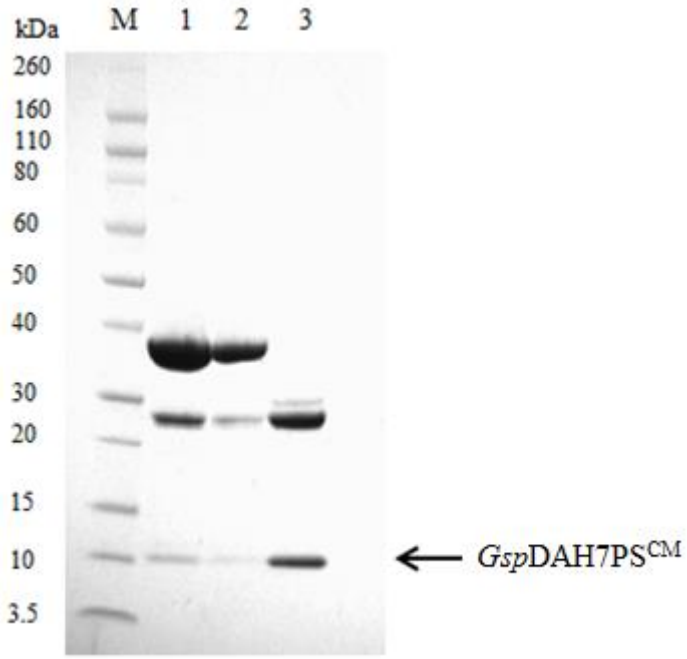


Figure 3.9: SDS-PAGE analysis for *Gsp*DAH7PS^{CM} after first GSTrap affinity column (lane 1), after desalting column (lane 2), after TEV cleavage (lane 3)

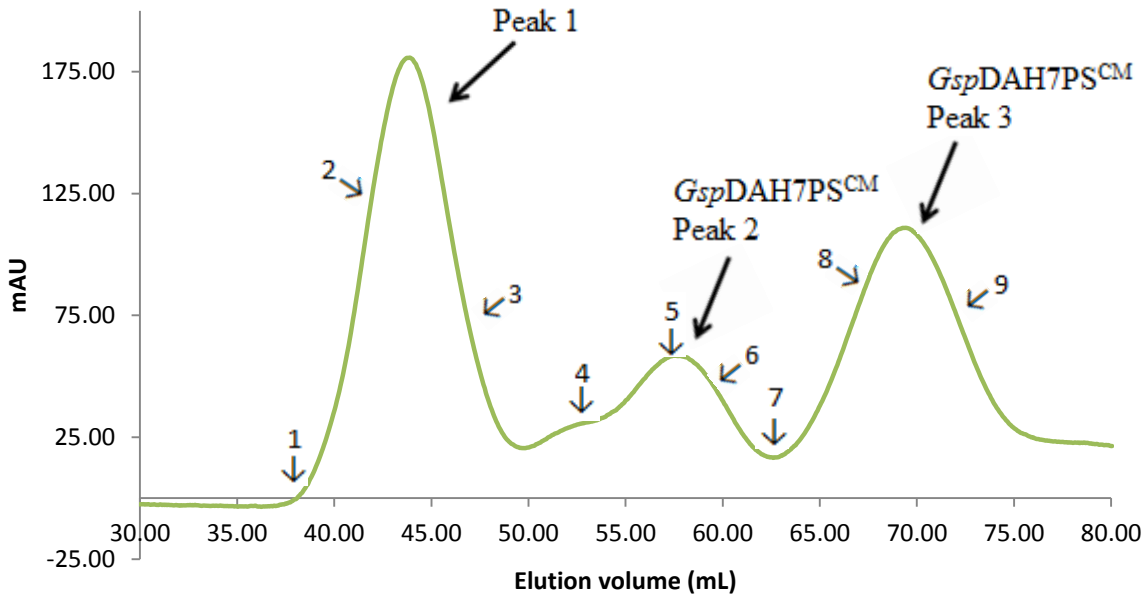


Figure 3.10: SEC trace of eluted *Gsp*DAH7PS^{CM} in SEC buffer (10 mM BTP, pH 7.4, 40 mM KCl, 200 µM PEP and 10 µM EDTA) showing three peaks corresponding to the eluted protein with the shoulder to the left of the second peak. Numbers 1 to 9 represent the lane number on the SDS-PAGE gel of where the fractions from the SEC run were taken (Figure 3.11).

SDS-PAGE analysis showed that only peaks two and three along with the shoulder contained *Gsp*DAH7PS^{CM}, with both peaks being active as determined by an enzyme activity assay. Peak one did not appear to contain any *Gsp*DAH7PS^{CM} and lacked CM activity (Figure 3.11). Peaks two and three were pooled separately avoiding peak 1 and the shoulder, concentrated down, put into aliquots, flash frozen and stored at -80 °C.

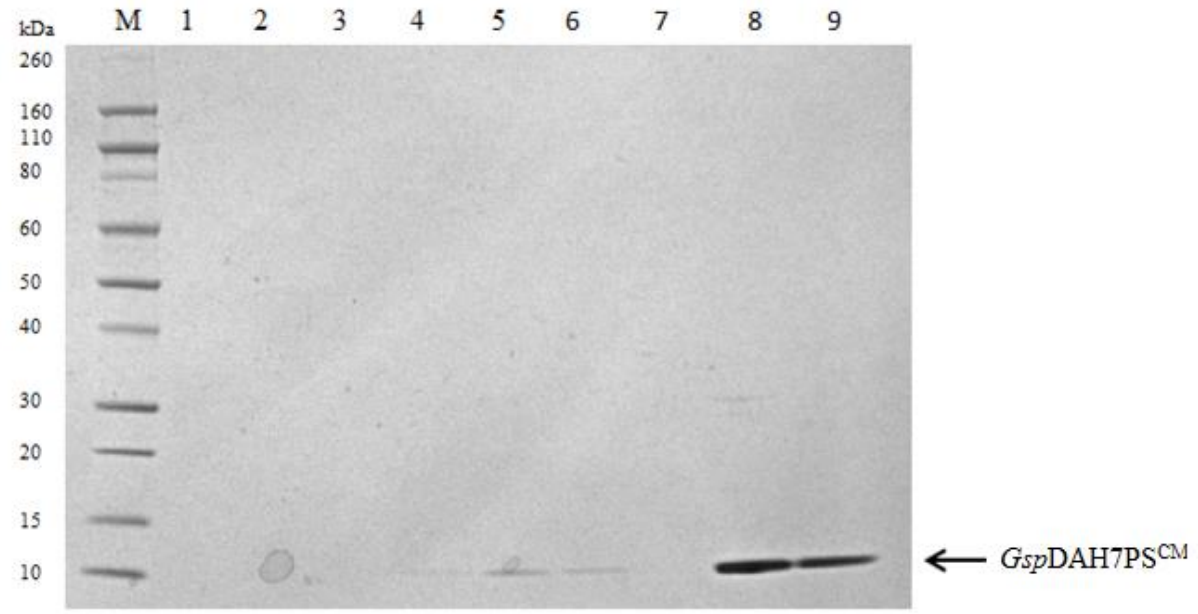


Figure 3.11: SDS-PAGE gel of final collected fractions after SEC of *Gsp*DAH7PS^{CM}. Lanes 1-3 show fractions collected from the first peak. Lanes 4-7 show fractions collected from the second peak. Lanes 8 and 9 show fractions collected from the third peak (Figure 3.10)

3.7 Mass spectrometry

The molecular masses of *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM} were determined using electrospray ionisation mass spectrometry to ensure the purified enzymes were of the expected mass calculated from the amino acid sequence using ProtPram. *Gsp*DAH7PS^{DAH7PS} had a mass of 30,095 Da and *Gsp*DAH7PS^{CM} had a mass of 11,149 Da, which is in agreement with the calculated masses of 30,095 Da and 11,149 Da respectively.

Once both variants were purified, different experiments were carried out in order to characterise each enzyme, the data obtained could then be used to compare to experimental data of the wild-type enzyme (Figure 3.12).

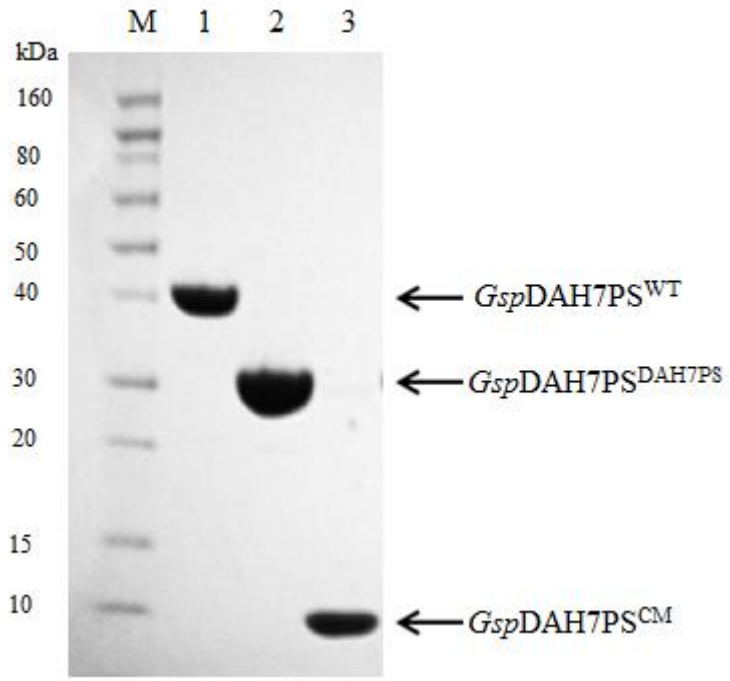


Figure 3.12: SDS-PAGE gel of purified *Gsp*DAH7PS^{WT} (lane1), *Gsp*DAH7PS^{DAH7PS} (Lane 2) and *Gsp*DAH7PS^{CM} (Lane 3) after SEC.

3.8 Protein secondary structure analysis

Circular dichroism (CD) spectra were obtained for *Gsp*DAH7PS^{DAH7PS} expected to comprise only the core barrel and *Gsp*DAH7PS^{CM}, which was expected to be α-helical. CD is an initial method of ensuring that the enzymes had folded properly and that no major conformational changes to the core barrel or the CM domain in the absence of the other component had occurred (Figure 3.14).

CD data for *Pfu*DAH7PS[65] was used to compare with the CD data obtained for *Gsp*DAH7PS^{DAH7PS} since both enzymes consist of only the core barrel (Figure 3.13).

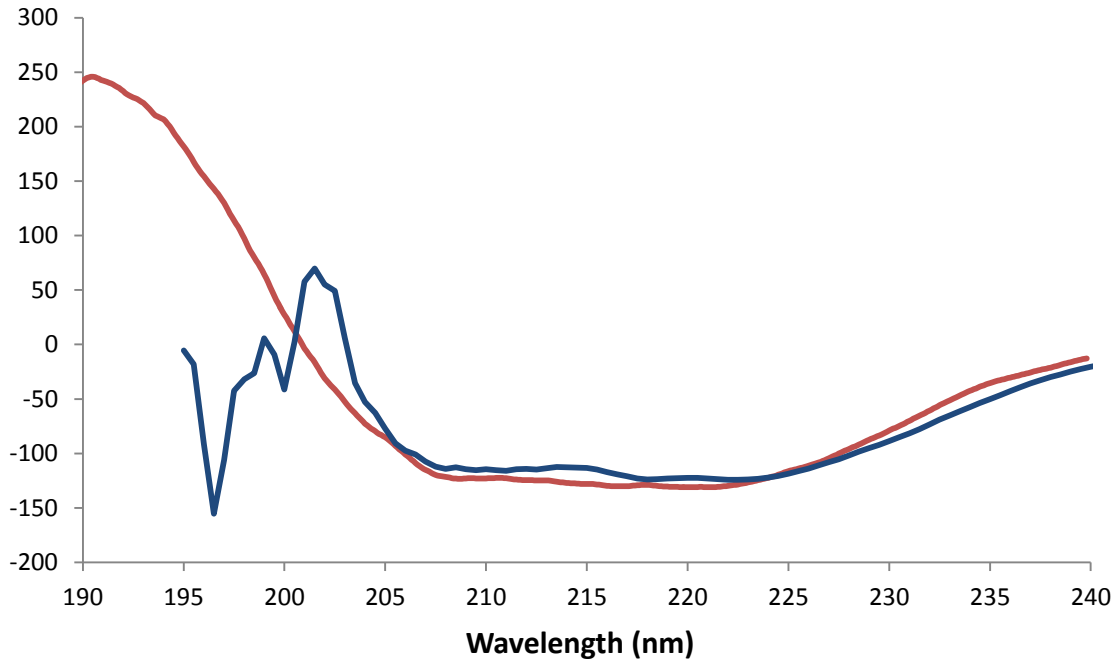


Figure 3.13: Circular dichroism spectra of *Gsp*DAH7PS^{DAH7PS} (red) and *Pfu*DAH7PS (dark blue). A protein concentration of 0.15 mg/mL for *Gsp*DAH7PS^{DAH7PS} and 0.01 mg/mL for *Pfu*DAH7PS was used exchanged into double-distilled water (pH 7.0).

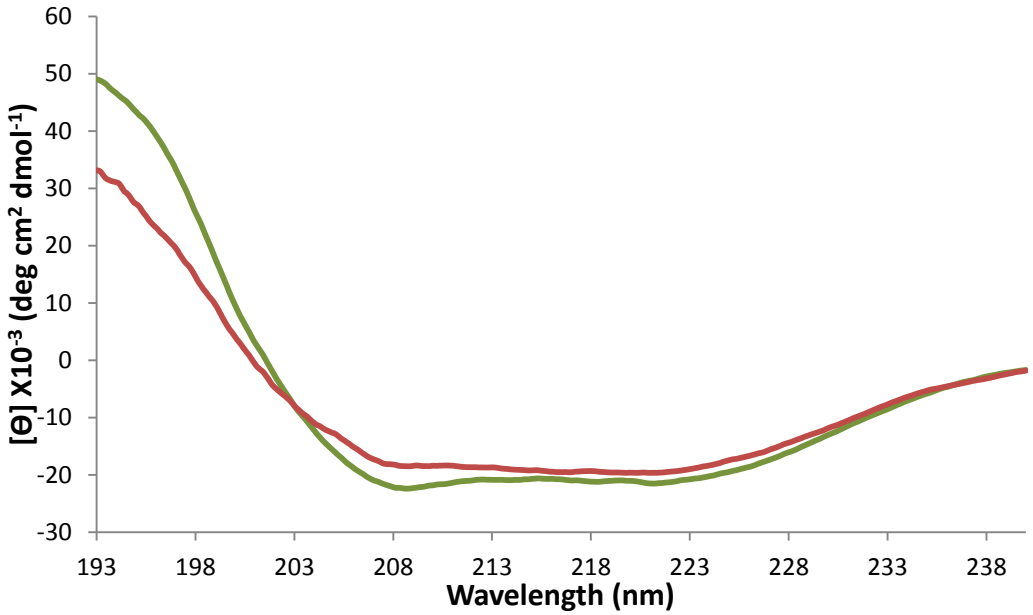
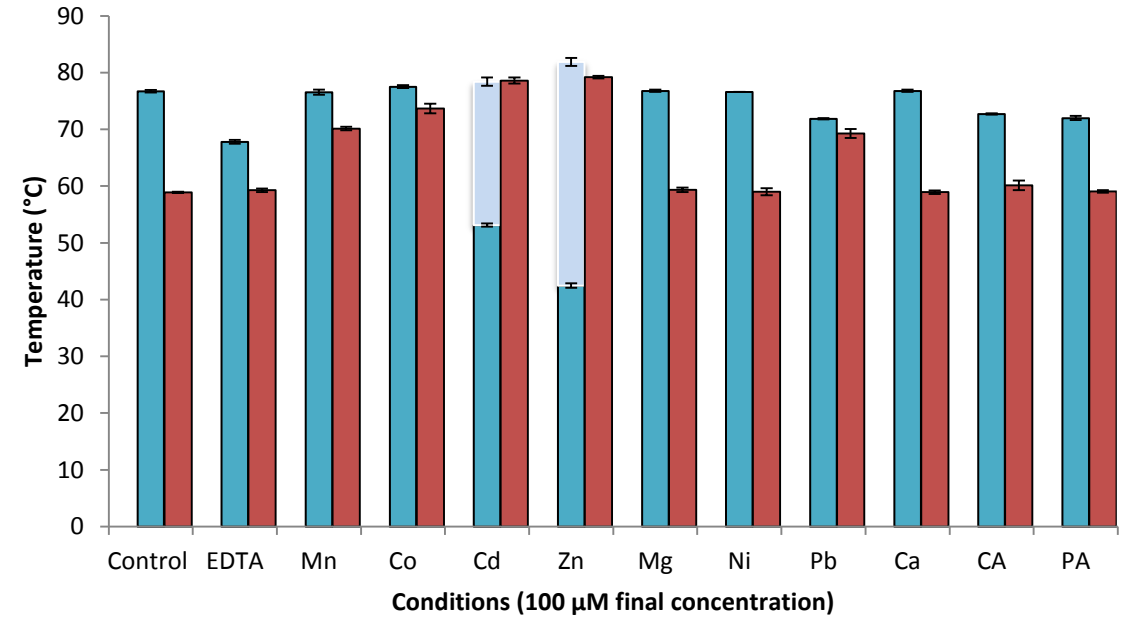


Figure 3.14: Circular dichroism spectra of *Gsp*DAH7PS^{DAH7PS} (red) and *Gsp*DAH7PS^{CM} (green). A protein concentration of 0.15 mg/mL was exchanged into double-distilled water (pH 7.0).

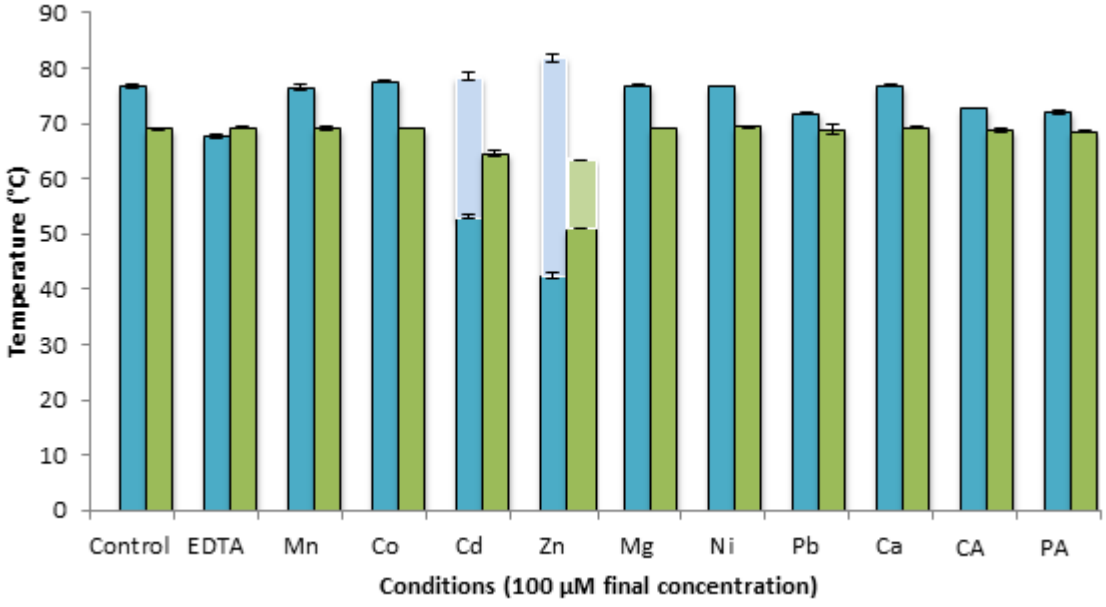
Although the CD spectrum of *Pfu*DAH7PS was of a poor quality, it showed similarities to the CD spectrum of *Gsp*DAH7PS^{WT} as anticipated. The trace for *Gsp*DAH7PS^{CM} resembles the expected traces for a helical protein. This data suggests that *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM} are properly folded and are suitable for further characterisation.

3.9 Differential scanning fluorimetry for *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM}

While DSF was carried out on *Gsp*DAH7PS^{WT}, it was important to determine the thermal stabilities of the separated regulatory CM domain and the catalytic DAH7PS domain under the same conditions used for the wild-type enzyme. Comparison of the data obtained for both mutants with the wild-type enzyme would reveal any effects on both domains when separated. The thermal stabilities of both *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM} were examined under different conditions (no metal-control, EDTA, Co^{2+}, Cd^{2+}, Zn^{2+}, Mg^{2+}, Ni^{2+}, Pb^{2+}, Ca^{2+}, CA, PA) using DSF.



(a)



(b)

Figure 3.15: DSF data showing T_m of (a) *Gsp*DAH7PS^{WT} (blue), *Gsp*DAH7PS^{DAH7PS} (red), (b) *Gsp*DAH7PS^{WT} (blue) and *Gsp*DAH7PS^{CM} (green) in the presence of different conditions (no metal-control), EDTA, Co^{2+}, Cd^{2+}, Zn^{2+}, Mg^{2+}, Ni^{2+}, Pb^{2+}, Ca^{2+}, CA, PA) at 100 µM with light blue bars indicating the secondary T_m for *Gsp*DAH7PS^{WT} and light green bars indicating the secondary T_m for *Gsp*DAH7PS^{CM}.

In the condition with no metal added (control) both *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM} are less stable than the wild-type enzyme with melting temperatures recorded at 58.9 ± 0.1 °C and 69 ± 0.2 °C respectively, in comparison to the temperature of 76.7 ± 0.2 °C observed for the wild-type enzyme. While the inclusion of EDTA has been shown to destabilise the wild-type enzyme, EDTA had no effect on stability for either *Gsp*DAH7PS^{DAH7PS} or *Gsp*DAH7PS^{CM} with T_m values of 59.3 ± 0.3 °C and 69.2 ± 0.2 °C respectively (Figure 3.15). Mn^{2+}, Co^{2+}, Cd^{2+}, Zn^{2+} or Pb^{2+} were the only additives that stabilised *Gsp*DAH7PS^{DAH7PS}, with Cd^{2+} and Zn^{2+} being the most stabilising, giving T_m values of 78.6 ± 0.5 °C and 79.2 ± 0.2 °C respectively.

For the wild-type enzyme, the presence of either Cd^{2+} or Zn^{2+} brought about two distinct rises in florescence, consistent with a two stage thermal denaturation response. This two stage process was not evident for *Gsp*DAH7PS^{DAH7PS} (Figure 3.15). Interestingly Cd^{2+} and Zn^{2+} were the only destabilising conditions for *Gsp*DAH7PS^{CM} , with Zn^{2+} causing the variant to also have a two stage thermal denaturation response with at 50.7 ± 0.1 °C for the first denaturation response, then 62.7 ± 0.1 °C for the second denaturation response. All other conditions did not appear to have any effect on the stability of the enzyme (Figure 3.15). The results suggest that Cd^{2+} and Zn^{2+} have a destabilising effect on the CM domain, yet are stabilising to the DAH7PS domain of *Gsp*DAH7PS^{WT}. This conclusion is supported by the fact that both metal ions are destabilising to *Gsp*DAH7PS^{CM} but stabilising to *Gsp*DAH7PS^{DAH7PS}. This would explain why the wild-type enzyme has a two stage thermal denaturation response to both metal ions.

3.10 Metal activation of *Gsp*DAH7PS^{DAH7PS}

A metal activation study carried out on *Gsp*DAH7PS^{WT} determined that the enzyme was most activated by Cd^{2+} (section 2.5). The same study was carried out for *Gsp*DAH7PS^{DAH7PS} in order to determine if the removal of the CM domain had any effect on the enzyme's ability to be activated by different metals.

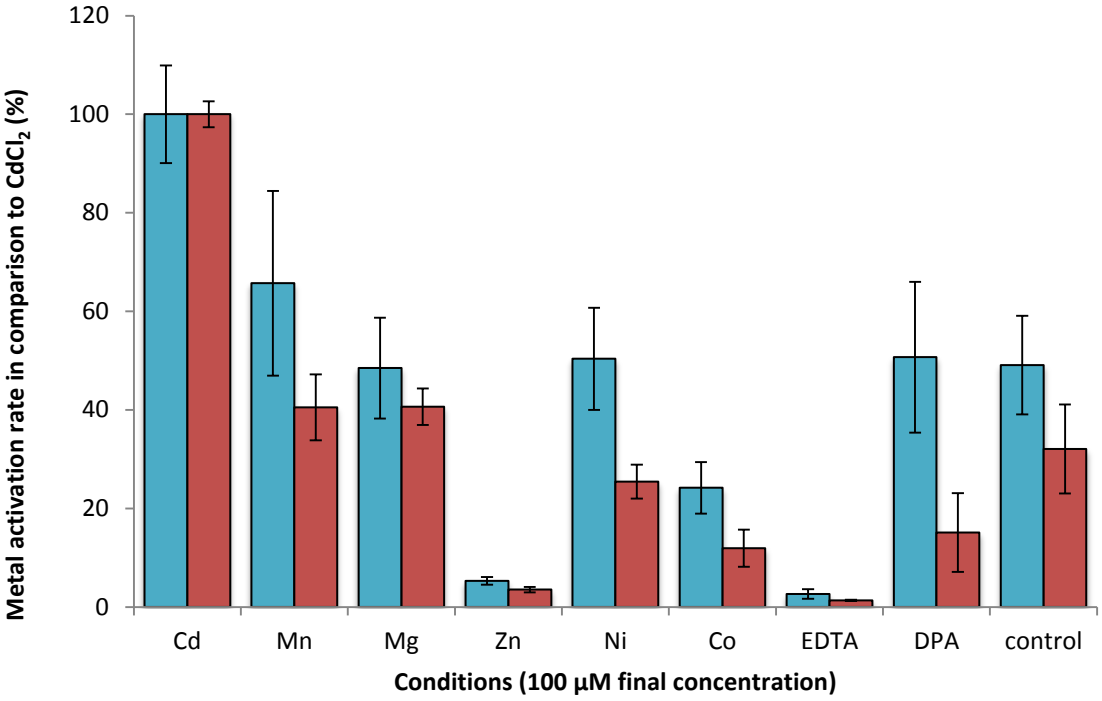


Figure 3.16: Metal activation graph for *Gsp*DAH7PS^{WT} (Blue) compared to *Gsp*DAH7PS^{DAH7PS} representing the different activation abilities of different metals in comparison to Cd^{2+} at a concentration of 100 µM. Also shown is the degree of inactivation for both enzymes in the presence of the metal chelators EDTA and DPA. Reactions were carried out at 60 °C, with 195 µM PEP and 197 µM E4P in 50 mM BTP buffer (pH 6.8).

As expected, it appears that the removal of the CM domain had little effect on the activation abilities of the different metals with Cd^{2+} still being the most activating metal and Zn^{2+} being the least activating metal ion in comparison to *Gsp*DAH7PS^{WT} (Figure 3.16).

3.11 Temperature dependency study of *Gsp*DAH7PS^{DAH7PS}

As for *Gsp*DAH7PS^{WT}, the optimal temperature for *Gsp*DAH7PS^{DAH7PS} activity was determined through a temperature dependency study using the same experimental methods (section 2.6). Substrates (PEP and E4P), metal (Cd^{2+}), buffer (50 mM BTP, pH adjusted to the required temperature) and enzyme concentrations were kept constant while varying the temperature of the reaction mixture. Once the reaction mixture in the cuvette had equilibrated to the required temperature, enzyme was added to the mixture and incubated for one minute before initiation of the reaction with E4P. The experiment was later

repeated with the enzyme incubated for longer period of five minutes, to determine the effects of a longer incubation time (Figure 3.17).

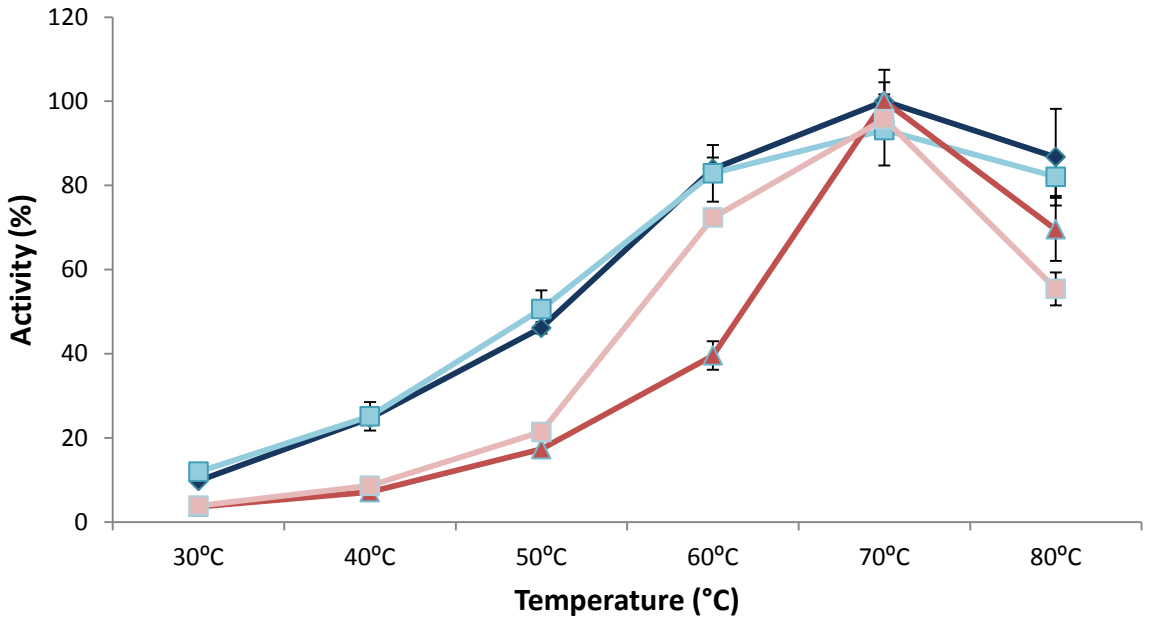


Figure 3.17: Temperature dependent graph for *Gsp*DAH7PS^{WT} and *Gsp*DAH7PS^{DAH7PS} showing enzyme activity at different temperatures. Substrate (325 µM PEP, 475 µM E4P for *Gsp*DAH7PS^{WT} and 337 µM PEP, 356 µM E4P for *Gsp*DAH7PS^{DAH7PS}), metal (100 µM Cd^{2+}), buffer (50mM BTP, pH 6.8) and enzyme (5 µL at 0.21 mg/mL for *Gsp*DAH7PS^{WT} and 3 µL at 0.14 mg/mL for *Gsp*DAH7PS^{DAH7PS}) concentrations were kept constant. Enzymes were incubated at desired temperature with the reaction mixture for either one minute (shown in dark blue for *Gsp*DAH7PS^{WT} and in red for *Gsp*DAH7PS^{DAH7PS}) or five minutes (shown in light blue for *Gsp*DAH7PS^{WT} and in light red for *Gsp*DAH7PS^{DAH7PS}).

The results show that just like the wild-type enzyme, *Gsp*DAH7PS^{DAH7PS} was most active at 70 °C and least active at 30 °C under the given conditions. For both enzymes, activity increases as temperature increases, up to temperatures as high as 70 °C, and then drops at 80 °C. This result suggest that both *Gsp*DAH7PS^{WT} and *Gsp*DAH7PS^{DAH7PS} may very well be denaturing at such high temperatures, with both enzymes determined to have a T_m below 80 °C according to DSF data (section 3.9), resulting in the decreased activity reading. Another possible explanation for the lost in activity at 80 °C could be a result of substrate degradation, where the substrate breaks down before reaching the active site of the catalytic DAH7PS domain. The different incubation times had no effect on the enzymatic

activity of *Gsp*DAH7PS^{DAH7PS} except at 60 °C, where the mutant was significantly more active when incubated for five minutes in comparison to one minute (Figure 3.17).

3.12 Temperature dependency study of *Gsp*DAH7PS^{CM}

The temperature dependency study was also carried out on *Gsp*DAH7PS^{CM} with the same method used for *Gsp*DAH7PS^{WT} where the substrate (CA), buffer (pH adjusted for experimental temperatures) and enzyme concentrations were kept constant while only varying the temperature (section 2.7). The enzyme was incubated (after equilibration was achieved) with the reaction mixture for either one minute or five minutes before initiation of the reaction with CA (Figure 3.18).

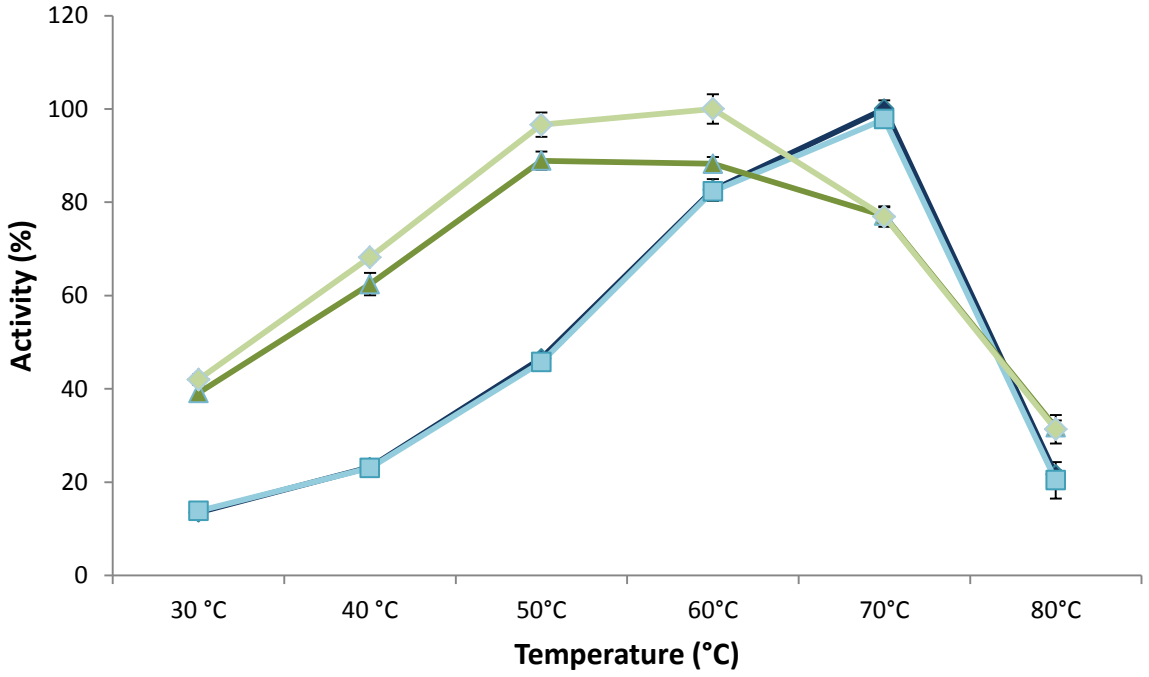


Figure 3.18: Temperature dependent graph for *Gsp*DAH7PS^{WT} and *Gsp*DAH7PS^{CM} showing enzyme activity at different temperatures (30 °C, 40 °C, 50 °C, 60 °C, 70 °C and 80 °C) with no added metal ions. Substrate (70 µM CA for *Gsp*DAH7PS^{WT} and 80 µM CA for *Gsp*DAH7PS^{CM}) buffer (50mM BTP, pH 6.8) and enzyme (3 µL at 12.97 mg/mL for *Gsp*DAH7PS^{WT} and 5 µL at 5.29 mg/mL for *Gsp*DAH7PS^{CM}) concentrations were kept constant. Enzyme was incubated at desired temperature with the reaction mixture for either one minute (shown in dark blue for *Gsp*DAH7PS^{WT} and in green for *Gsp*DAH7PS^{CM}) or five minutes (shown in light blue for *Gsp*DAH7PS^{WT} and in light green for *Gsp*DAH7PS^{CM}).

The results show that both *Gsp*DAH7PS^{WT} and *Gsp*DAH7PS^{CM} CM enzyme activity increases as temperature increases. The wild-type enzyme is most active at 70 °C, and then the activity dramatically drops at 80 °C, possibly due to substrate degradation. This is unlike the truncation mutant that is most active at only 60 °C, with enzymatic activity reduced at 70 °C and decreased to a greater extent at 80 °C (Figure 3.18). This result is consistent with the DSF data for *Gsp*DAH7PS^{CM} that shows that under similar conditions the enzyme has a T_m of 69 ± 0.2 °C, significantly lower than the wild-type enzyme which under the same conditions has a T_m of 76.7 ± 0.2 °C (section 3.9). This result suggests that the truncation mutant at 70 °C is slowly denaturing and therefore giving the reduced activity reading.

Incubating the wild-type enzyme with the reaction mixture for either one or five minutes did not have an effect on the activity of the enzyme. In the case of the truncation mutant, it appears that incubating the enzyme with the reaction mixture at a temperature of choice meant that the enzyme was slightly more active (up to temperatures of 60 °C), at five minutes of incubation, than incubating it for one minute with the reaction mixture (Figure 3.18). This suggests that *Gsp*DAH7PS^{CM} requires a longer incubation time to reach full activity.

3.13 Michaelis-Menten kinetics of *Gsp*DAH7PS^{DAH7PS}

A comparison of kinetic parameters between *Gsp*DAH7PS^{WT} and *Gsp*DAH7PS^{DAH7PS} was required in order to find out just how much of an effect the removal of the CM domain had on the catalytic domain's ability to mediate its reaction. The steady state kinetic parameters for *Gsp*DAH7PS^{DAH7PS} were measured using the same spectrophotometric assay at 60 °C (Figure 3.19). The apparent K_m values for PEP and E4P were found to be 98 ± 11 µM and 62 ± 5 µM respectively, with a calculated k_{cat} value of 63 ± 5 s^{-1} (Table 3.1).

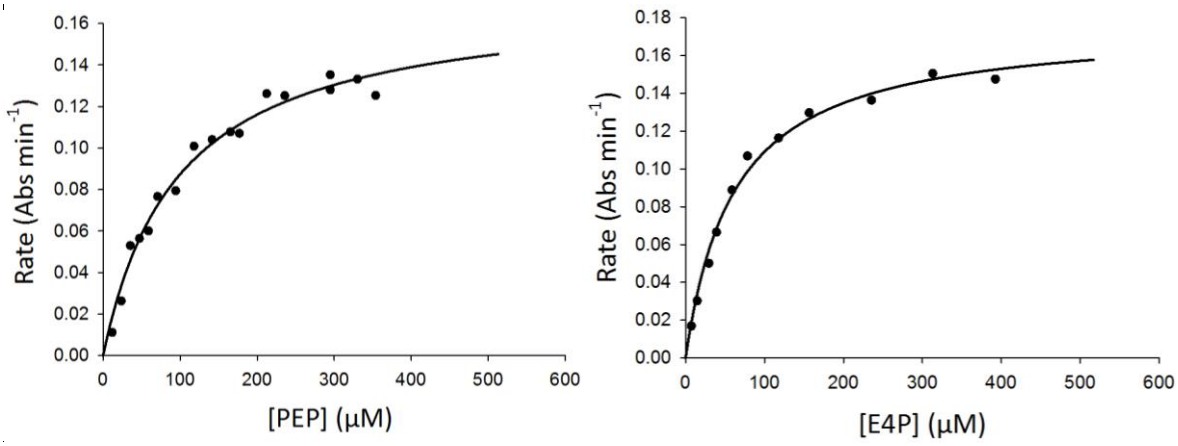


Figure 3.19: Michaelis-Menten plots obtained for *Gsp*DAH7PS^{DAH7PS} showing the apparent PEP (left) K_m with a value of 98 ± 11 µM and E4P (right) K_m with a value of 62 ± 5 µM. A k_{cat} value of 63 ± 5 s^{-1} was obtained. Standard assays were carried out at 60 °C in 50 mM BTP and 100 µM Cd^{2+}.

Compared to the full-length wild-type protein, *Gsp*DAH7PS^{DAH7PS} has a similar K_m^{PEP} value, with a slightly lower K_m^{E4P} value. *Gsp*DAH7PS^{DAH7PS} is more active than *Gsp*DAH7PS^{WT} with a k_{cat} value that is roughly 1.4 fold higher than the k_{cat} value calculated for *Gsp*DAH7PS^{WT} (45 ± 4 s^{-1}).

Table 3.1: Kinetic parameters for the enzymes *Gsp*DAH7PS^{WT} and *Gsp*DAH7PS^{DAH7PS}

Enzyme	K_m^{PEP} (µM)	K_m^{E4P} (µM)	k_{cat} (s^{-1})	k_{cat}/K_m^{PEP} (µM.s^{-1})	k_{cat}/K_m^{E4P} (µM.s^{-1})
***Gsp*DAH7PS^{WT}**	87 ± 9	95 ± 8	45 ± 4	0.52	0.47
***Gsp*DAH7PS^{DAH7PS}**	98 ± 11	62 ± 5	63 ± 5	0.64	1.0

3.14 Michaelis-Menten kinetics of *Gsp*DAH7PS^{CM}

The steady state kinetic parameters for *Gsp*DAH7PS^{CM} were measured by tracking the consumption of CA at 274 nm (Figure 3.20). The continuous spectrophotometric assay for the enzyme was carried out at 50 °C

Initial velocities for the catalytic CM activity of *Gsp*DAH7PS^{CM} were measured by initiating the equilibrated reaction mixture with varying concentrations of CA. The apparent K_m value *Gsp*DAH7PS^{CM} was found to be 450 ± 42 µM, with a calculated k_{cat} value of 5.5 ± 0.52 s^{-1} (Table 3.2).

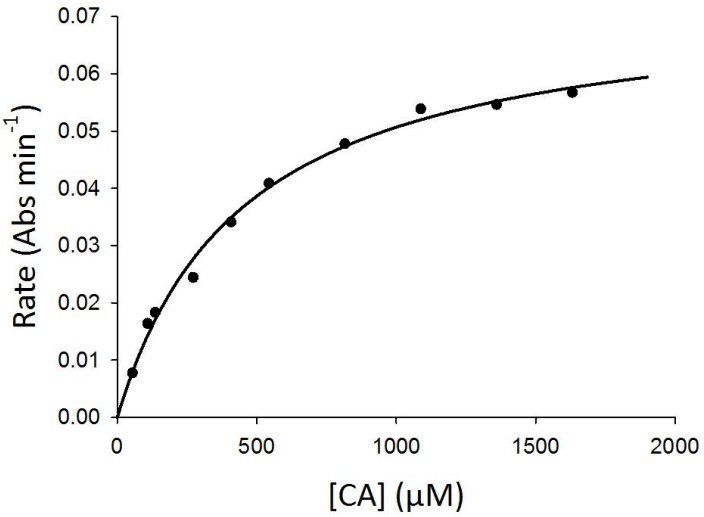


Figure 3.20: Michaelis-Menten plot obtained for *Gsp*DAH7PS^{CM} showing the apparent CA K_m with a value of 450 ± 42 µM . A k_{cat} value of 5.5 ± 0.52 s^{-1} was obtained for *Gsp*DAH7PS^{CM}. Standard assays were carried out at 50 °C in 50 mM BTP (pH 6.8) without the addition of any metal ions.

Compared to the full-length wild-type enzyme, *Gsp*DAH7PS^{CM} has a K_m^{CA} value that is significantly higher. *Gsp*DAH7PS^{CM} is more active, at maximal activity, than the wild-type enzyme with a k_{cat} value that is roughly 2.9 fold higher. Analysis of the k_{cat}/K_m^{CA} values suggests that *Gsp*DAH7PS^{WT} is the better catalyst at 2.1x10^{-2} µM.s^{-1} in comparison to *Gsp*DAH7PS^{CM} with a value of 1.2x10^{-2} µM. s^{-1}.

Table 3.2: Kinetic parameters for the enzymes *Gsp*DAH7PS^{WT} and *Gsp*DAH7PS^{CM}

Enzyme	K_m^{CA} (µM)	k_{cat} (s^{-1})	k_{cat}/K_m^{CA} (µM.s^{-1})
***Gsp*DAH7PS^{WT}**	88 ± 6	1.9 ± 0.13	2.1x10^{-2}
***Gsp*DAH7PS^{CM}**	450 ± 42	5.5 ± 0.52	1.2x10^{-2}

3.15 Feedback inhibition of *Gsp*DAH7PS^{DAH7PS}

An inhibition study was carried out on *Gsp*DAH7PS^{DAH7PS} to investigate what effect the removal of the CM domain has on the response of the enzyme to PA (Figure 3.21). DAH7PS domain activity was tested using a continuous spectrophotometric assay at 60 °C, where the consumption of PEP was monitored at 232 nm. Substrates (PEP, E4P), metal (Cd^{2+}), and

enzyme concentrations were kept constant in 50 mM BTP buffer (pH 6.8), while the concentration of PA was varied.

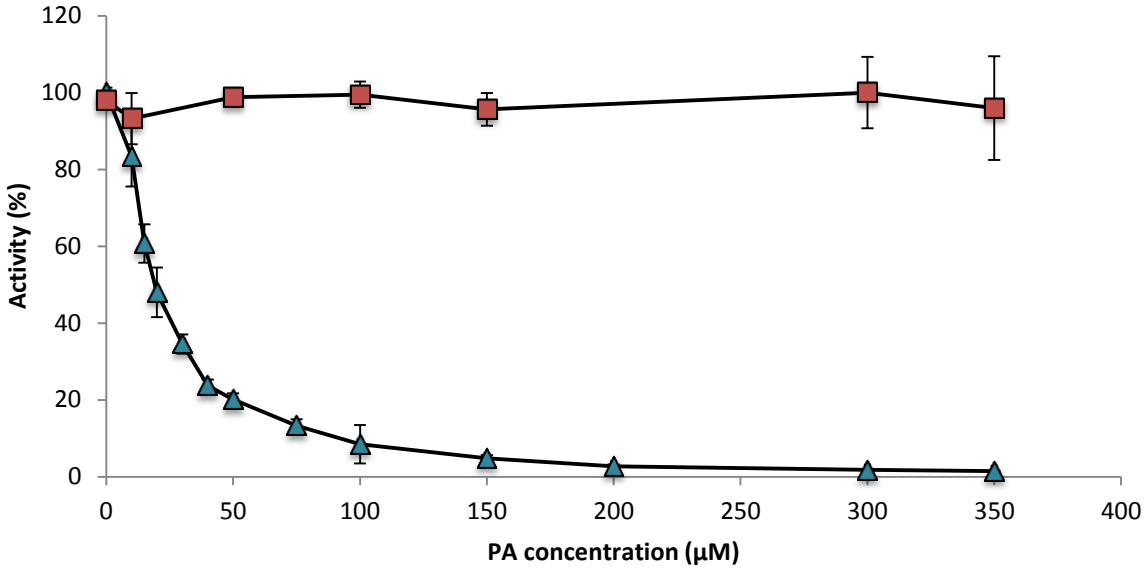


Figure 3.21: Response of *Gsp*DAH7PS^{WT} (blue) and *Gsp*DAH7PS^{DAH7PS} (red) activity to increasing concentrations of PA. Standard assays were carried out at 60 °C, with 121 µM PEP and 123 µM E4P in 50 mM BTP buffer (pH 6.8). The data points were measured in triplicate with error bars indicating the standard deviation.

Results show no reduction in enzymatic activity for *Gsp*DAH7PS^{DAH7PS} in the presence of PA at any concentration. This is in contrast to the wild-type enzyme, which was completely inhibited by PA and shows that the CM extension to the barrel is necessary for inhibition by PA.

3.16 Analytical SEC of *Gsp*DAH7PS^{DAH7PS}

Another technique used to examine the structure of *Gsp*DAH7PS^{DAH7PS} was analytical size exclusion chromatography (SEC) (gel filtration) (Figure 3.22). This technique determines molecular masses of proteins by using a column that separates proteins based on their size, with larger proteins eluting first followed by the smaller proteins. In the case of *Gsp*DAH7PS^{DAH7PS}, this technique was used to determine if the enzyme exists as a single or multiple species in solution. By loading different enzyme concentrations onto the column,

the technique also determines whether the quaternary structure of the enzyme is concentration dependent. The analytical SEC column was calibrated using protein standards with known molecular weights, to allow estimation of the molecular weight of the protein being eluted, based on elution volume. The molecular weight of eluted *Gsp*DAH7PS^{DAH7PS} was determined based on a standard curve of the log of molecular weight against elution volume/void volume (V_e/V_o) of protein standards.

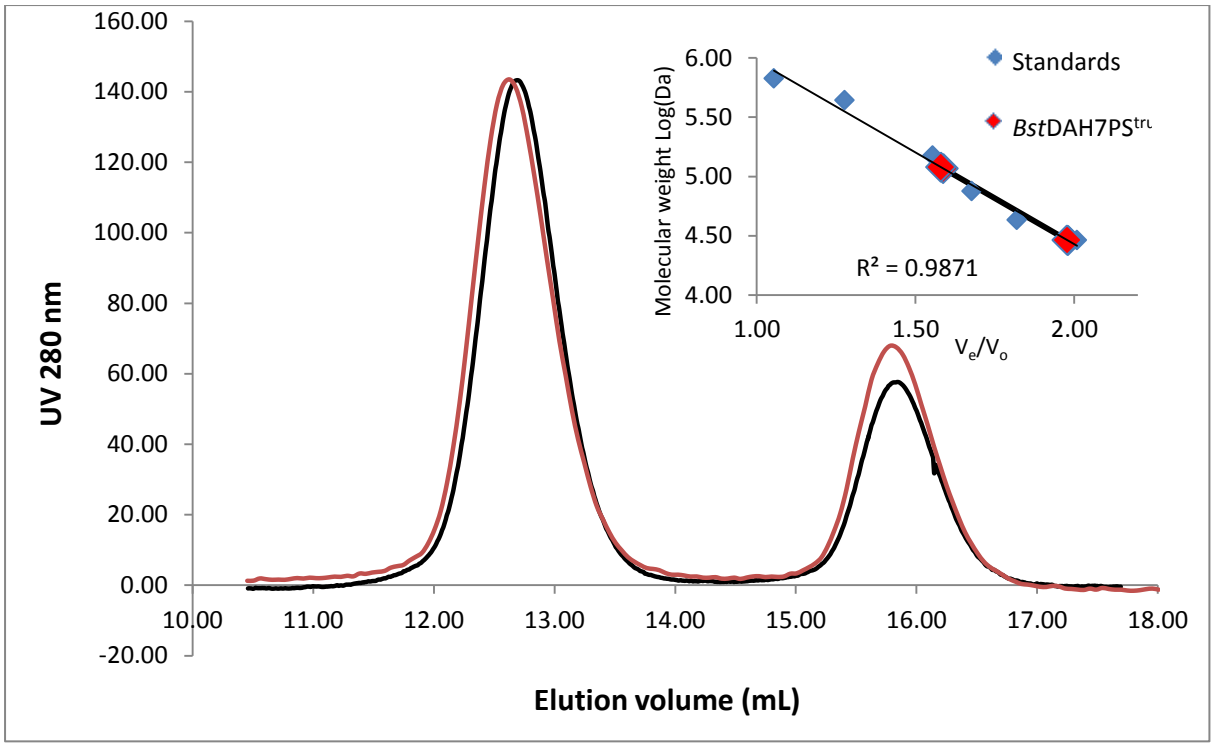


Figure 3.22: Analytical SEC traces (normalised) and standard curve of log molecular weight versus elution volume/void volume for *Gsp*DAH7PS^{DAH7PS} at 1 mg/ml (black trace) and 0.5 mg/ml (red trace). Injection volume was 500 µL for each sample.

Two different concentrations (1 mg/mL, 0.5 mg/mL) of *Gsp*DAH7PS^{DAH7PS} were injected onto the column with the enzyme eluting in two peaks for both runs. The enzyme eluted in the first peak corresponded to the molecular weights of 117 kDa and 120 kDa for the first run and the second run respectively. The enzyme eluted in the second peak corresponded to the molecular weight of 28.9 kDa and 29.1 kDa for the first run and the second run respectively. This indicates that the enzyme exists both as a monomer and a tetramer in solution. There is

a small shift in oligomeric state favouring tetramer formation as the enzyme concentration is increased.

3.17 Analytical ultracentrifugation of *Gsp*DAH7PS^{DAH7PS}

Gel filtration traces of *Gsp*DAH7PS^{DAH7PS} show the enzyme elutes in two peaks. SDS-PAGE analysis determined that both peaks contain *Gsp*DAH7PS^{DAH7PS}. Activity assays showed that both of the peaks were active. Analytical SEC data indicated that the enzyme exists in equilibrium between a tetrameric and monomeric state. To further investigate the quaternary assembly of *Gsp*DAH7PS^{DAH7PS} in solution, an analytical ultracentrifugation (AUC) experiment was run on different concentrations of the enzyme (Figure 3.23).

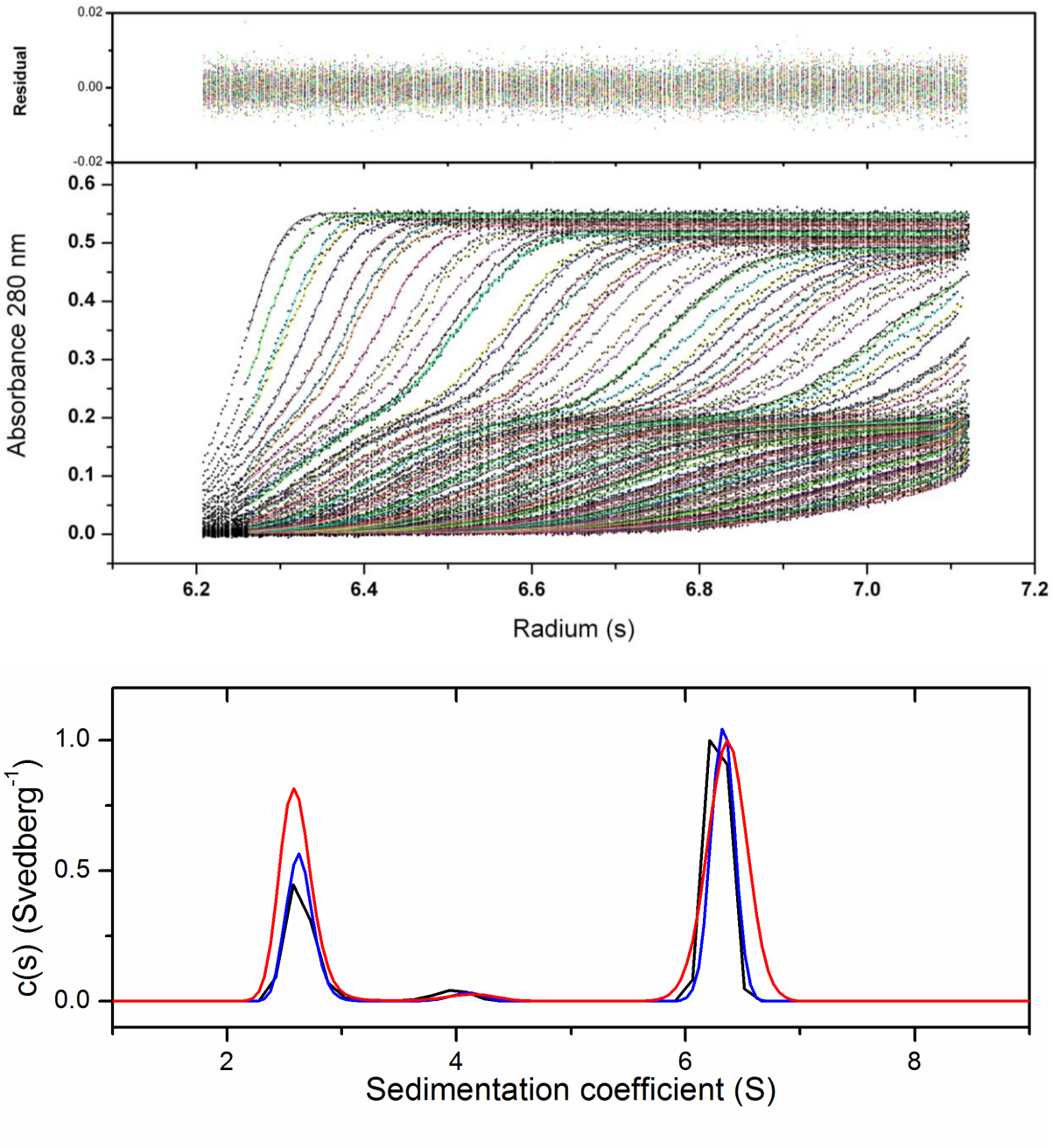


Figure 3.23: Sedimentation velocity analysis of *Gsp*DAH7PS^{DAH7PS} at 1.8 mg/ml (black line), 1 mg/ml (blue line) and 0.5 mg/ml (red line). Experiments were carried in SEC buffer (pH 7.4 at 20 °C). Data was collected at 50,000 rpm at a wavelength of 280 nm. (A) the best fit of the experimental data. (B) the distribution of the species.

Stock solutions of *Gsp*DAH7PS^{DAH7PS} were buffer exchanged into freshly made SEC buffer (10 mM BTP, pH 7.4, 40 mM KCl, 200 µM PEP and 10 µM EDTA) and diluted to three different concentrations (1.8 mg/mL, 1 mg/mL and 0.5 mg/mL). Using AUC each sample was run along with a blank containing an exact buffer match to ensure a precise profile of the enzyme in solution is obtained.

The absorbance versus radial position data show the presence of three sedimentation boundaries for *Gsp*DAH7PS^{DAH7PS}, suggesting that under the experiemental conditions, the enzyme exists as three oligomeric species in solution. A best fit of the experimental data showed the presence of three species that had sedimentation coefficients of 2.6 S, 4.1 S and 6.4 S. Fitting the sedimentation velocity data, correlated to calculated masses of 37 kDa, 72 kDa and 130 kDa for *Gsp*DAH7PS^{DAH7PS}. This is consistent with the molecular masses for the monomer (30 kDa), dimer (60 kDa) and tetramer (120 kDa) of *Gsp*DAH7PS^{DAH7PS}.

Sedimentation equilibrium experiments at different concentrations were carried out to ensure that equilibrium between the three oligomeric states was being observed. Normalising the tetrameric sedimentation coefficient boundaries obtained to one for all concentrations showed that as the concentration of enzyme increased, tetramer formation was favoured with a reduced monomer formation.

3.18 Summary

The genes encoding for *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM} were both amplified from the plasmid containing the *Gsp*DAH7PS^{WT} gene and ligated into the vectors of choice, which were then transformed into the desired host cells. Both truncated variants were expressed and purified successfully by using the appropriate columns, with size exclusion chromatography being the final step in the purification procedure. Mass spectrometry determined that both enzymes were of the expected mass.

As for *Gsp*DAH7PS^{WT}, DSF results showed that different metal ions in comparison to EDTA had different stabilising effects on *Gsp*DAH7PS^{DAH7PS}. With no added metal, both *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM} were less stable in comparison to the wild-type enzyme. Cd^{2+} and Zn^{2+}, the two metal ions that brought about two distinct rises in florescence consistent with a two stage thermal denaturation response for the wild-type enzyme, were the most stabilising metal ions for *Gsp*DAH7PS^{DAH7PS}, and most destabilising to *Gsp*DAH7PS^{CM}.

*Gsp*DAH7PS^{DAH7PS} was found to be metal dependent without much change to the activation abilities of the metal ions to the truncated mutant in comparison to the wild-type enzyme, with Cd^{2+} being the most activating metal ion for the enzyme similar to that observed for *Gsp*DAH7PS^{WT}.

Both truncated variant activities were found to be affected by temperature where *Gsp*DAH7PS^{DAH7PS} gave similar results to the wild-type enzymes DAH7PS activity, with both enzymes being most active at 70 °C. *Gsp*DAH7PS^{CM} gave a slightly different result with the highest enzymatic activity observed at 60 °C, unlike the wild-type enzyme with maximal CM activity at 70 °C.

Michaelis-Menten kinetics obtained through kinetic assays determined that *Gsp*DAH7PS^{DAH7PS} has a similar K_m^{PEP} value (98 ± 11 µM), with a slightly lower K_m^{E4P} (62 ± 5µM) value in comparison to the wild-type enzyme that had K_m values of 87 ± 9 µM and 95 ± 8 µM for PEP and E4P respectively. *Gsp*DAH7PS^{DAH7PS} was also found to be more active than *Gsp*DAH7PS^{WT} with a k_{cat} value (63 ± 5 s^{-1}) that is roughly 1.4 fold higher than the k_{cat} value calculated for *Gsp*DAH7PS^{WT} (45 ± 4 s^{-1}). *Gsp*DAH7PS^{CM} had a significantly higher K_m^{CA} value (450 ± 42 µM) Compared to *Gsp*DAH7PS^{WT} with a K_m^{CA} value of 88 ± 6 µM. *Gsp*DAH7PS^{CM}

was found to be more active that the wild-type protein with a k_{cat} value that is roughly 2.9 fold higher, yet is a less efficient catalyst at $1.2x10^{-2}$ µM. s^{-1} in comparison to *Gsp*DAH7PS^{WT} with a value of $2.1x10^{-2}$ µM.s^{-1}.

While inhibition studies determined that *Gsp*DAH7PS^{WT} was completely inhibited by PA, *Gsp*DAH7PS^{DAH7PS} showed no signs of inhibition in the presence of the inhibitor. It was therefore determined that the CM domain plays a key role in the regulation of the wild-type enzyme.

Analytical SEC and AUC data demonstrated that *Gsp*DAH7PS^{DAH7PS} exists in multiple oligomeric states in solution with the data suggesting that the truncated mutant is in equilibrium between the two states.

Chapter 4

Summary of thesis and future directions

This thesis describes the characterisation of the type Iβ bifunctional DAH7PS fusion protein, *Gsp*DAH7PS^{WT}, and the further characterisation carried out on the independently expressed truncated domains.

*Gsp*DAH7PS^{WT} demonstrated both DAH7PS and CM activity, with the catalytic DAH7PS domain requiring a divalent metal ion for catalytic activation. PA completely inhibited the DAH7PS activity of *Gsp*DAH7PS^{WT}, and according to AUC, causes the wild-type enzyme to favour a tetrameric oligomeric state in solution rather than display the dimer-tetramer equilibrium that was observed for *Gsp*DAH7PS^{WT} in the absence of PA.

4.1 The regulatory and catalytic domains of *Gsp*DAH7PS^{WT} were successfully separated and expressed independently and both domains retained activity.

Analysis of the sequence of *Gsp*DAH7PS^{WT} suggested that the protein consists of two functional proteins, a DAH7PS and a CM. The bifunctionality of this protein was unequivocally confirmed through the kinetic characterisation of the protein. Separating the DAH7PS and CM domains of *Gsp*DAH7PS^{WT} and expressing these independently revealed that both domains, *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM}, retained activity on their own. These observations demonstrated that both domains are capable of functioning independently of one another without a requirement to be tethered.

*Gsp*DAH7PS^{DAH7PS} not only retained activity, but was determined to be more catalytically active than the wild-type enzyme according to kinetic assays, with a k_{cat} value 1.4 fold higher than that recorded for the wild-type enzyme. This increase in activity was also observed for the type Iβ *Tma*DAH7PS, where removal of the fused regulatory FL domain resulted in a more catalytically active enzyme, with a k_{cat} value 1.6 fold higher than the wild-type

protein.[16] Interestingly, the removal of the CM domains found fused to *Bsu*DAH7PS and *Pgi*DAH7PS hardly had an effect on the activity for the isolated DAH7PS of *Bsu*DAH7PS, and resulted in a *Pgi*DAH7PS domain that had reduced catalytic activity, with a k_{cat} value 2 fold less in comparison to the full length enzyme.[12]

*Bsu*DAH7PS, like *Gsp*DAH7PS^{WT}, is another fusion enzyme with an N-terminal CM domain. A previous study characterising *Bsu*DAH7PS suggested that this enzyme was not a true bifunctional protein, as the CM domain fused to the catalytic domain had very poor activity.[12] *B. subtilis* contains another gene encoding for a second highly active AroH CM domain responsible for the conversion of CA to PA.[60] This CM enzyme serves the organism's primary needs of PA production, rendering the catalytic activity of the CM fused to the DAH7PS redundant. This further supported the idea that *Bsu*DAH7PS had acquired this CM fusion for the purpose of regulation, rather than for bifunctionality.

Geobacillus sp also contains another gene encoding for a second intracellular AroH CM domain, however in the case of *Gsp*DAH7PS^{WT}, it should be noted that the enzyme appears to be bifunctional as the fused AroQ CM domain in this protein has a much higher activity in comparison to the AroQ CM domain fused to *Bsu*DAH7PS. It would be of great interest to characterise the AroH CM domain of *Geobacillus sp* and compare its catalytic efficiency to that obtained with the AroQ CM domain of *Gsp*DAH7PS^{WT} in this study. It would be interesting to determine whether both the AroH and AroQ CM proteins available to *Geobacillus sp* have comparable kinetic parameters. It is conceivable that the AroH CM domain of *Geobacillus sp* may have some sort of functional role that the AroQ CM domain of *Gsp*DAH7PS^{WT} cannot fulfil. One way to determine this would be to genetically modify *Geobacillus sp* to encode for the fused AroQ CM and not the AroH CM, then observe what impact this has on the viability and the properties of the organism.

4.2 The removal of the regulatory CM domain left the catalytic *Gsp*DAH7PS^{DAH7PS} domain activity unaffected by PA, suggesting that the CM domain plays a key role for the allosteric regulation of *Gsp*DAH7PS^{WT}

PA was shown to completely inhibit the catalytic activity of *Gsp*DAH7PS^{WT} unlike *Gsp*DAH7PS^{DAH7PS}, which showed no signs of inhibition in the presence of PA at any concentration. This suggests that the CM domain of *Gsp*DAH7PS^{WT} is necessary for inhibiting DAH7PS activity. Clearly while there are benefits for *Geobacillus sp* in having a regulated DAH7PS enzyme, its presence incurs the penalty of reduced catalytic activity. As for *Gsp*DAH7PS^{DAH7PS}, the truncated catalytic domains of *Bsu*DAH7PS and *Pgi*DAH7PS did not demonstrate any sign of inhibition in the presence of inhibitors.[12] This was also the case for *Tma*DAH7PS, which was no longer regulated by pathway end products upon removal of the FL domain, implying that the FL domain was also required for the regulation of the catalytic domain.[16] This was confirmed when a comparison between ligand bound and non-ligand bound crystal structures of *Tma*DAH7PS revealed the enzymes mechanism of inhibition, whereby the regulatory FL domains from diagonally opposite subunits interact to form a binding site for Tyr, resulting in physical capping of the active site and inhibition of DAH7PS activity.

Elucidating the mechanism of inhibition for *Gsp*DAH7PS^{WT} requires further experimental work. However, speculation can be made as to how inhibition of the enzyme occurs by utilising the data already obtained, and using the structure of the related *Lmo*DAH7PS to complement this data.[61] *Lmo*DAH7PS consists of a catalytic DAH7PS domain fused to an N-terminal CM domain and shares 75 % primary sequence identity to *Gsp*DAH7PS^{WT}. While no functional characterisation has been carried out on *Lmo*DAH7PS, the structural characterisation demonstrates its similarity to *Gsp*DAH7PS^{WT}. The core of the tetrameric state of both enzymes consists of four $(\beta/\alpha)_8$ barrels that contain the catalytic DAH7PS machinery. These core barrel domains alternate in their orientations, where diagonally related subunits face the same side of the tetramer. Each DAH7PS catalytic subunit is connected to an N-terminal CM domain through a linker region where two CM domains

emerging from diagonally related DAH7PS subunits that face the same side of the tetramer come together and interact with one another (Figure 4.1).

The crystal structure of *Gsp*DAH7PS^{WT} was solved with PA-bound[65] and displays a distinct conformation to the non-ligand bound crystal structure of *Lmo*DAH7PS. Formation of CM dimers loosely associated with DAH7PS subunits is believed to be a key part in the inhibition of *Lmo*DAH7PS and *Gsp*DAH7PS^{WT}. It is believed that upon binding of PA to the CM dimers, a conformational change occurs where loosely associated CM dimers close in and cap the active site of the DAH7PS domains, blocking substrate access and therefore inhibiting the catalytic domain. The conformational change that occurs is easily achieved due to the highly flexible linker connecting the catalytic and regulatory domains, allowing for the capping of DAH7PS active sites possible without restrictions. This mechanism of inhibition is further supported when comparing the crystal structure of PA-bound *Gsp*DAH7PS^{WT} with a non-PA bound model of *Gsp*DAH7PS^{WT} based on the *Lmo*DAH7PS PA-bound crystal structure. (Figure 4.1). The non-PA bound model of *Gsp*DAH7PS^{WT} shows that the CM dimers are loosely associated with the tetramer in an open conformation, unlike the PA-bound structure which clearly shows the CM dimers closely associated with the tetramer in a closed form to function as an allosteric cap.

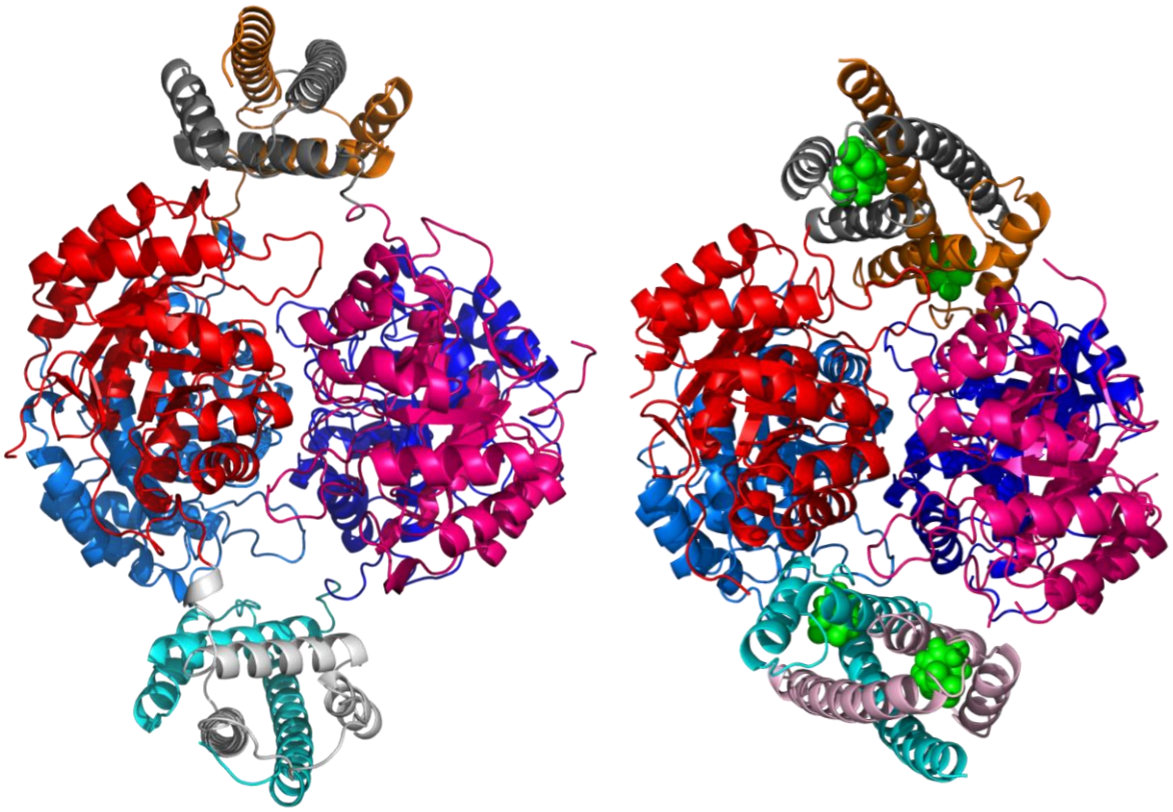


Figure 4.1: non-ligand bound (left) and ligand bound (right) tetrameric structures of *Gsp*DAH7PS^{WT}, with different colours representing different subunits. For the structure on the right, each CM subunit has PA-bound shown as green spheres. The structure on the left is a homology model based on the structure of *Lmo*DAH7PS (PDB code 3TFC).

The conformational change that *Gsp*DAH7PS^{WT} undergoes upon PA binding is not as extensive as that of *Tma*DAH7PS, where the regulatory FL domains must come together to form a binding site for the inhibitor leading to the capping of the active site.[16] In *Gsp*DAH7PS^{WT} the regulatory CM domains are already interlocked and in close proximity to the active site, ready to cap the active site of the catalytic domain. This could be of advantage to *Gsp*DAH7PS^{WT}, where such a mechanism would lead to a quicker and more precise inhibition. Indeed it has been clearly demonstrated that *Gsp*DAH7PS^{WT} is highly sensitive to PA according to the inhibition study carried out.

4.3 The removal of the CM domain disrupted the tetrameric oligomeric state of *Gsp*DAH7PS^{WT}, thus the catalytic CM regulatory domain supports protein structure in addition to allosteric regulation.

SEC traces of *Gsp*DAH7PS^{WT} displayed a peak with a shoulder, both determined to contain *Gsp*DAH7PS^{WT} according to SDS-PAGE analysis. Further experimental work on the wild-type enzyme using AUC, determined that the enzyme exists in equilibrium between a dimeric and a tetrameric oligomeric state. AUC was also performed on *Gsp*DAH7PS^{DAH7PS}, since SEC traces for this enzyme also displayed two peaks, with both peaks determined to contain *Gsp*DAH7PS^{DAH7PS} according to SDS-PAGE analysis. AUC results determined that the truncated variant mostly exists in equilibrium between a monomeric and tetrameric state, with a small amount of dimer observed. A similar result was also observed by analytical SEC, which also determined that the enzyme is in equilibrium between a monomeric and tetrameric oligomeric state. Therefore it is clear that the removal of the ancillary CM domain from the catalytic DAH7PS domain of *Gsp*DAH7PS^{WT} causes a disturbance in the quaternary structure, leading to monomer formation. This suggests that the CM domain not only regulates DAH7PS activity, but also supports the quaternary structure of the enzyme. While truncating the CM domain from *Gsp*DAH7PS^{WT} disrupted the quaternary structure, the resulting catalytic activity of the enzyme was not compromised, and instead it was established that *Gsp*DAH7PS^{DAH7PS} is relatively more active in comparison to *Gsp*DAH7PS^{WT}. This observation suggests that tetramer formation is not necessary for DAH7PS activity.

Preliminary results suggest that the reason for the observed increase in catalytic activity of *Gsp*DAH7PS^{DAH7PS} in comparison to *Gsp*DAH7PS^{WT} is due to the removal of the CM domains that would normally be in close proximity to the active sites of the DAH7PS domains. Even if inhibition of *Gsp*DAH7PS^{WT} requires the CM domains to cap the DAH7PS active sites of the tetramer, it is a possibility that in the open more extended confirmation the loosely associated CM dimers may well impede catalysis and access of substrates and product release. This may account for the observation that *Gsp*DAH7PS^{DAH7PS} is more active than the wild-type protein, since each DAH7PS domain is more accessible to substrates by not having a CM subunit in close proximity to the active site. According to small angle X-ray scattering

(SAXS) data, the regulatory FL domains of *Tma*DAH7PS are rapidly opening and closing even in the absence of an inhibitor.[16] This alternating conformational change could explain why the removal of the regulatory domain produces a variant more catalytically active, where the FL domain no longer sterically hinders substrate access to the active site of the DAH7PS domain.

While the above explanation for the increased catalytic activity observed agrees with the results obtained for *Gsp*DAH7PS^{WT} and *Tma*DAH7PS, another explanation could be that the increase in catalytic activity observed for *Gsp*DAH7PS^{DAH7PS} is related to the disrupted quaternary structure resulting from the CM domain truncation. This idea is supported by similar results obtained for *Tma*DAH7PS, where experimental data suggested that the removal of the FL domain caused the enzyme to adopt a dimeric quaternary structure in the crystal structure instead of the tetrameric structure adopted by the wild-type enzyme (16 this thesis). SAXS data did however show that the *Tma*DAH7PS variant was a tetramer, suggesting that the enzyme can populate a different quaternary structure and does not necessarily have to be tetrameric. Another finding that greatly supports this explanation comes from a recent study carried out on *Pfu*DAH7PS, where the residue Ile, located in the hydrophobic interface of the enzyme, was substituted for the residue Asp.[70] The resulting mutant was found to exist in equilibrium between a dimeric and tetrameric quaternary state, unlike the wild-type enzyme which only exists as a single tetrameric species in solution. Interestingly, disturbing the quaternary structure of the wild-type enzyme resulted in a variant that was 4 fold catalytically more active.

If removing the CM domain of *Gsp*DAH7PS^{WT} causes the enzyme to fall apart but consequently have an increased catalytic activity, why would *Geobacillus sp* evolve such a fusion enzyme at the expense of reduced catalytic activity? It has been clearly demonstrated that *Gsp*DAH7PS^{DAH7PS} has no sensitivity to PA, showing that the CM domain is required for regulation. The mechanism of inhibition of *Gsp*DAH7PS^{WT} requires the enzyme to adopt a tetrameric oligomeric state, clearly shown by the crystal structure of the wild-type enzyme and by AUC, where *Gsp*DAH7PS^{WT} tetramer formation increased significantly in the presence of PA. Enzyme regulation provides a great advantage to an organism, as it provides the ability to control pathway output in response to metabolic demand. A DAH7PS enzyme

possessing the trait of allosteric regulation may be of great importance to *Geobacillus sp*. It is therefore a possibility that this thermophilic bacterium advantageously evolved the regulated *Gsp*DAH7PS^{WT}, even at the cost of diminished activity. It would be interesting to see what would occur, if *Geobacillus sp* was genetically modified to encode for *Gsp*DAH7PS^{DAH7PS} rather than *Gsp*DAH7PS^{WT} in order to determine what impact this has on the viability and the properties of the organism.

The quaternary structure of *Gsp*DAH7PS^{WT} appears to be beneficial not for regulation, but also for increased thermostability. *Geobacillus sp* is a thermophillic organism thriving in environmental conditions with extreme temperatures.[64] Possessing enzymes that are able to withstand the high temperatures this organism normally inhabits is of importance. DSF data demonstrated that *Gsp*DAH7PS^{WT} was more stable at evaluated temperatures in comparison to *Gsp*DAH7PS^{DAH7PS} in all but one of the conditions tested, where both enzymes were equally stable. The study mentioned above where the tetrameric structure of *Pfu*DAH7PS was disrupted by a single point mutation, also determined that the resulting mutant was significantly less stable at elevated temperatures in comparison to the wild-type enzyme.[70] This further supports the idea that in addition to regulation, the quaternary structure of *Gsp*DAH7PS^{WT} promotes added thermostability. So whereas *Gsp*DAH7PS^{DAH7PS} is more catalytically active in comparison to *Gsp*DAH7PS^{WT}, it is less thermostable.

Further investigations using AUC and DSF experiments on other type Iβ DAH7PS enzymes such as *Bsu*DAH7PS and *Pgi*DAH7PS, along with their truncated variants that lack a regulatory CM domain may prove to be relevant. Characterisation work to help determine the quaternary structure and thermostability of the full length enzymes, along with the truncated catalytic domains, could provide a better understanding of why DAH7PS enzymes display an increase in catalytic activity as a result of having their regulatory domains removed. Since it has been shown that the removal of the CM domains found fused to *Bsu*DAH7PS and *Pgi*DAH7PS hardly had an effect on the activity for the DAH7PS truncated mutant of *Bsu*DAH7PS, and resulted in a *Pgi*DAH7PS truncated mutant with diminished activity, would AUC data reveal a lack of disturbance in quaternary structure due to the truncation? In addition, would DSF data of the full length enzymes and their respective

truncated mutants demonstrate that they have the same thermostability due to a lack of disturbance in quaternary structure?

It is unclear why *Gsp*DAH7PS^{WT} is in equilibrium between a dimeric and tetrameric oligomeric state rather than a tetrameric state only. This finding does seem unusual, especially since the enzyme requires the tetrameric oligomeric state for inhibition to occur. It is worth noting however that other AUC data carried on *Gsp*DAH7PS^{WT} in the presence of higher salt concentrations identified the presence of tetramer only. Further investigation is required into the quaternary structure of *Gsp*DAH7PS^{WT}, where the enzyme is accessed with AUC under different environmental conditions ranging in pH, temperature and salt concentration. It is possible that the cellular conditions of *Geobacillus sp* promote the formation of tetramer only, which would be favourable for regulation and stability of the enzyme, especially under extreme temperatures.

4.4 Future experiments

Now that it has been determined that the CM domain of *Gsp*DAH7PS^{WT} is required for regulation, it would be interesting to determine if it is possible to transfer the CM domain of *Gsp*DAH7PS^{WT}, and fuse it to the catalytic barrel of a DAH7PS enzyme from a different source. It has been previously shown that transfer of the FL domain of *Tma*DAH7PS to the unregulated *Pfu*DAH7PS, resulted in a chimeric enzyme that was regulated by Tyr, and to a lesser extent by Phe.[50] Would transferring the CM domain of *Gsp*DAH7PS^{WT} to *Pfu*DAH7PS form a chimeric enzyme that is inhibited by PA? What would occur if the same experiment was carried out but instead, fusing the CM domain to *Pfu*DAH7PS previously mentioned that contains the single point mutation that disrupts the homotetramer?

The reverse experiment could also be carried out where the catalytic domain of *Gsp*DAH7PS^{WT} is fused to the regulatory domain of another type Iβ DAH7PS enzyme such as the FL domain *Tma*DAH7PS. Would this chimeric protein have comparable kinetic parameters to *Gsp*DAH7PS^{WT}?

Conclusion

Understanding the mechanism of regulation in DAH7PS enzymes is a very important part of unravelling how these enzymes evolved and the reasons behind it. *Gsp*DAH7PS^{WT} evolved to be regulated through the acquisition of an ancillary CM domain. The results obtained through the characterisation of *Gsp*DAH7PS^{WT} provided some insight into the enzymes mechanism of regulation and in turn, contributed towards a better understanding of the diverse methods of regulation found in the DAH7PS enzyme family.

Chapter 5

Experimental procedures

5.1 General methods

5.1.1 Amino acid sequence alignments

Multiple amino acid sequence alignments were generated using ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

5.1.2 Protein structure figures

All protein structure figures were created using the PyMOL™ Molecular Graphics System (version 1.3, Schrödinger LLC)

5.1.3 Purified water

All water was purified by passage through a Millipore Milli-Q water system. This is referred to as Milli-Q water throughout this thesis. All Milli-Q water used for molecular biology was autoclaved prior to use.

5.1.4 pH measurement

The pH of all solutions was measured at room temperature or at kinetic assay temperatures unless otherwise stated using a Mettler Toledo SevenCompact™ 220 pH probe. The pH of solutions was adjusted using 1 M or 10 M HCl and 1 M or 10 M NaOH.

5.1.5 Removal of metal ions from solution

Metal ions were removed from solution by treatment with Chelex® 100 resin (Bio-Rad). The resin was added and left to stir overnight before the resin was removed by filtration. The pH of all solutions was re-adjusted after Chelex® treatment.

5.1.6 Antibiotic stocks

Ampicillin (100 mg/mL), kanamycin (100 mg/mL) and spectinomycin (10 mg/mL) stocks were all made up using milliQ water unlike chloramphenicol (25 mg/mL) which was made up in analytical grade ethanol. All solutions were filtered and then aliquoted into sterile microtubes and stored at 80 °C.

5.1.7 Centrifugation

All centrifugation was carried out using the following:

- Sorvall Evolution RC (Thermo Scientific)
- Heraeus Multifuge 1 S-R (Thermo Scientific)
- Heraeus Pico 17 (Thermo Scientific)
- Minispin® Centrifuge (Eppendorf)

5.1.8 Growing *E. coli* cells

For every 1 L culture, a 50 mL preculture containing the appropriate antibiotic was inoculated from the glycerol stock and grown at 37 °C overnight. Once grown, cells were transferred to the baffled flasks containing lysogeny broth (LB) solution and correct antibiotics where the main culture was grown at 37 °C with shaking at 180 RPM until an OD_{600} of 0.4-0.6 was reached.

5.1.9 Cell induction

Proteins expression was achieved with the addition of IPTG (0.5 mM final concentration) to the cells that had reached the desired OD_{600}. Cultures were then incubated at 37 °C for four

hours for the expression of *Gsp*DAH7PS^{WT}, or incubated at 23 °C overnight for the expression of the *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM} variants. Cultures incubated at 23 °C were left to sit at room temperature for twenty min before the addition of IPTG to ensure protein expression at the desired temperature.

5.1.10 Harvesting of cells

Cells were harvested (four hours after induction for *Gsp*DAH7PS^{WT} and one night after induction for *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM}) centrifugation (14000 *g*, 4 °C, twenty five min). Once cells pellets were obtained, they were stored at -80 °C until needed.

5.1.11 Cell lysis

Cells were lysed by sonication using an Omni- Ruptor 4000 Ultrasonic Homogenizer (Omni international) with either a titanium 3/8 inch or 5/32 inch probe. 20 ml lysis buffer containing half a tablet of protease inhibitor cocktail tablet, EDTA (SIGMAFAST^{TM}) dissolved was used to resuspend the cell pellets of the His-tagged proteins. The suspended cells were placed in a beaker surrounded with ice and sonicated with 4-5 multiples of five min, at 80 % power for the 3/8 inch probe and 40 % power for the 5/32 inch probe with the pulser set at 35.

The lysis buffer used for *Gsp*DAH7PS^{WT} was composed of 50 mM Bis-tris propane (BTP), 200 mM KCl, 200 µM PEP and 2mM dithiothreitol (DTT) pH 7.3.

The lysis buffer used for *Gsp*DAH7PS^{DAH7PS} was composed of 20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole, pH 7.4 (HisTrap HP Binding buffer).

The lysis buffer used for *Gsp*DAH7PS^{CM} was composed of 140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3 (GSTrap binding buffer)

Proteins were separated from lysed cells by centrifugation (14000 *g*, 4 °C, twenty five min, or 40,000 g, 4 °C, twenty five min)

5.1.12 Measuring protein concentration

A Nanodrop ND-1000 spectrophotometer (Thermo Scientific) was used to measure the concentrations of all proteins by adopting Beers Law. The buffer that the protein was stored in was used to blank the Nanodrop, and 2 µL of the purified protein along with its calculated extinction coefficient value (http://web.expasy.org/protpram/) were used to calculate the concentration.

5.1.13 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed using pre-cast gels using NuPAGE® 10 % Bis-tris Gel 1.0 mm × 12-well pre-cast gels (Invitrogen) in NuPAGE® MOPS SDS Running Buffer (Invitrogen). Electrophoresis was performed using a XCell SureLock™ Electrophoresis Cell (Invitrogen), at 200 V for fifty min. Samples were mixed with NuPAGE® LDS Sample Buffer (4×) and DTT, and boiled before loading.

5.1.14 Buffer exchange and concentration of protein

Proteins were concentrated during and after purification using either a 20 mL 10 000 Da molecular weight cut-off (MWCO) device for *Gsp*DAH7PS^{WT} and *Gsp*DAH7PS^{DAH7PS} or a 20 mL 5000 Da MWCO device for *Gsp*DAH7PS^{CM} (Sartorius Stedim Biotech). All filtration devices were washed before use with 2-3 equivalent volumes of Milli-Q water. Protein solutions were buffer-exchanged by dilution with the desired buffer and repeatedly concentrated.

5.1.15 Determination of substrate concentration

The concentrations of PEP and E4P were determined using the standard assay system. Used by Schoner and Herrmann,[71] as modified by Schofield et al.[40] for use at 60 °C. The consumption of PEP was monitored at 232 nm (ϵ of 2.6 × 10^3 M^{-1} cm^{-1} at 60 °C) using a Varian Cary 100 UV Visible Spectrophotometer. In the case of E4P, a reaction mixture consisting of approximately 300 µM PEP, 100 µM Cd^{2+}, enzyme and 50 mM BTP buffer (pH 6.8) is made up before being initiated with the addition of 4µL of E4P stock solution

(unknown concentration) to give a final concentration of 1 mL. The reaction is allowed to go to completion and the difference in absorbance before initiation with E4P and the absorbance at completion of the reaction is measured. To convert absorbance into concentration the Beer-Lambert Law is applied (A = ε c l, where l = 1cm and ε of 2.6 x 10^3 M^{-1} cm^{-1} at 60 °C).

5.1.16 Storage of enzymes

Before storage, all purified enzymes were split into aliquots of 50 µL, snap frozen in liquid nitrogen and stored at −80 °C

5.1.17 Measurement of molecular weight

The masses of purified proteins were measured by electrospray ionisation using Bruker maXis™ 3G (Bruker Daltonics). Proteins were diluted from stored concentrations to 1 mg mL^{-1} with Milli-Q water.

5.1.18 Assays performed at 60 °C

Reaction mixtures containing 50mM BTP buffer (adjusted to have a pH of 6.8 at 60 °C), $CdCl_2$ and PEP were incubated for six min at 60 °C before the addition of enzyme. The reaction mixture was further incubated for one to five min depending on the experiment being carried out, before initiation the reaction with the addition of E4P. Triplicate readings of all measurements were recorded and averaged.

5.2 Methods for chapter two

5.2.1 Differential scanning fluorimetry

Melt temperatures of all the proteins were determined in the presence of different additives by DSF, using an iCycler iQ5™ Multicolour Real-Time PCR Detection System (Bio-Rad). Triplicate protein samples were added with mixing to buffer (containing additives) and SYPRO orange dye (Life Technologies) in a 96-well microplate, which was then sealed. The

melt proceeded in 0.2 °C increments from 20 °C to 95 °C, with a 20 s dwell time after each temperature rise. Measurements of the fluorescence were made at the end of each dwell time.

For each sample (three replicates and blank) 100 µL of buffer (10 mM BTP, pH 7.4, 40mM KCl, 200 µM) including any additives(s) was prepared, to which 5 µL of 250x SYPRO orange dye was added. In the blank well for the sample, 21 µL of the solution was added to 4 µL of water or buffer. Then, 16 µL of protein (at 1 mg/mL) was added to the remaining condition solution before 25 µL was dispensed into three replicate wells. In all experiments, each additive was at a final concentration of 100 µM.

For analysis an Excel spread sheet with custom VBA-scripted macro was created to toggle and dynamically display melt and derivative curves for each sample. The melt temperatures were calculated as the temperature of maximum infliction of the melting curve after subtracting the reading of a blank well containing buffer and dye but lacking protein.

5.2.2 Metal activation

The metal activation profile of *Gsp*DAH7PS^{WT} was determined using standard assays with 195 µM PEP and 197 µM E4P in 50 mM BTP buffer Chelex®-treated and filtered, pH 6.8 at 60 °C and 100 µM Cd^{2+}, Mn^{2+}, Mg^{2+}, Zn^{2+}, Ni^{2+}, Co^{2+}, EDTA, DPA or no condition (control). The enzyme (5 µL at 0.23 mg/mL) was incubated and left to equilibrate with the reaction mixture for 5 min at 60 °C prior to activation with E4P, to a final volume of 1 mL Triplicate assays were performed and averaged.

5.2.3 Temperature dependency study

The temperature dependency profile of *Gsp*DAH7PS^{WT} was determined using standard assays. Substrate concentrations were kept constant with 325 µM PEP and 475 µM E4P in 50 mM BTP buffer Chelex®-treated and filtered, pH 6.8, at the desired temperature with 100 µM Cd^{2+}. The reaction mixture was left to equilibrate at the chosen temperature. At each temperature (30 °C, 40 °C, 50 °C, 60 °C, 70 °C and 80 °C), the enzyme, 5 µL at 0.21 mg/mL ($0.26x10^{-2}$ µM in cuvette) was then incubated and left to equilibrate with the reaction

mixture for either one or five min prior to activation with E4P, to a final volume of 1 mL. Triplicate assays were performed.

5.2.4 Temperature dependency study (chorismate mutase activity)

The temperature dependency profile of *Gsp*DAH7PS^{WT} observing catalytic CM activity was determined using standard assays. The substrate concentration was kept constant with 70 µM CA in 50 mM BTP buffer Chelex®-treated and filtered, pH 6.8 at the desired temperature. The reaction mixture was left to equilibrate at the chosen temperature. At each temperature (30 °C, 40 °C, 50 °C, 60 °C, 70 °C and 80 °C), the enzyme, 3 µL at 12.97 mg/mL (0.94 µM in cuvette) was then incubated and left to equilibrate with the reaction mixture for either one or five min prior to activation with CA, to a final volume of 1 mL. Triplicate assays were performed.

5.2.5 Michaelis-Menten kinetics

The assays performed to determine the kinetic parameters for *Gsp*DAH7PS^{WT} were carried out in 1 cm path length quartz cuvettes. Standard reaction mixtures contained 100 µM $Cd^{2+,}$ 5 µL at 0.232 mg/mL ($2.8x10^{-2}$ µM in cuvette) (K_m^{PEP}), 5 µL at 0.288 mg/mL ($3.5x10^{-2}$ µM in cuvette) (K_m^{E4P}) enzyme in 50 mM BTP buffer (pH 6.8), giving a total volume of 1 mL. To determine K_m^{E4P}, the PEP concentration was fixed at 338 µM while the E4P concentration was varied (12 µM to 384 µM). For determination of K_m^{PEP}, E4P concentration was fixed at 329 µM while the PEP concentration was varied (27 µM to 410 µM).

5.2.6 Michaelis-Menten kinetics (chorismate mutase activity)

The assays for determining the kinetic parameters for *Gsp*DAH7PS^{WT} were carried out in 1 cm path length quartz cuvettes. Standard reaction mixtures contained 2 µL 12.97 mg/mL (0.63 µM in cuvette) enzyme in 50mM BTP buffer (pH 6.8), giving a total volume of 1 mL. To determine the K_m^{CA} different concentrations of CA were used varying form (7 µM to 298 µM)

5.2.7 Feedback inhibition

The inhibition profile of *Gsp*DAH7PS^{WT} was determined using standard assays. Varying concentrations of PA were added to make a 1 mL solution consisting of Chelex®-treated and filtered 50 mM BTP buffer (pH 6.8 at 60 °C), Cd^{2+} (100 µM), PEP (121 µM) and enzyme (10 µM at 0.11 mg/mL). The reaction mixture was incubated at 60 °C for ten min to equilibrate prior to initiating the reaction by addition of E4P (123 µM). All reactions were performed in triplicate and averaged.

5.2.8 Analytical ultracentrifugation

Sedimentation velocity and sedimentation equilibrium experiments were performed in a Beckman Coulter Model XL-I analytical ultracentrifuge equipped with UV/Vis scanning optics. Reference buffer solution (10 mM BTP, pH 7.4, 40 mM KCl, 200 µM PEP and 10 µM EDTA) and sample solutions were loaded into 12 mm double-sector cells with quartz windows. The cells were then mounted in an An-60 Ti 4-hole rotor. *Gsp*DAH7PS^{WT} samples at a concentration of 1.8 mg/mL, 1 mg/mL and 0.5 mg/mL (380 µL) and reference (400 µL) were centrifuged at 50,000 rpm at 20 °C and absorbance data were collected in continuous mode at 278 nm without averaging. The sample that was run in the presence of PA was run at a concentration of 1 mg/ml in the presence of 500 µM PA, with a reference buffer solution of 10 mM BTP, pH 7.4, 40 mM KCl, 200 µM PEP, 10 µM EDTA and 500 µM PA.

Data were fitted to a continuous size distribution model[72] using the program SEDFIT. The partial specific volume (v) of the sample (0.7463 mL g^{-1}), buffer density (1.0004 g mL^{-1}) and buffer viscosity (1.004 cP) were calculated using the program SEDNTERP.[73]

5.3 Methods for chapter three

5.3.1 Amplification of *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM} ORF

The plasmid containing the gene coding for *Gsp*DAH7PS^{WT} was used as the template for cloning out the ORF of *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM} variants. This was achieved by

designing the appropriate primers and the use of the iProof^{TM} High-Fidelity DNA polymerase PCR kit (BioRad).

Nested PCR was the technique used to clone both truncation variants which required two PCR rounds, with the second round adding extensions encoding the recombination sites to the amplified genes which allows the gene products to be inserted into the desired entry vector.

*Gsp*DAH7PS^{CM} primers (Round 1)

Forward primer: **5'-GGC AGC GGC GCG ATG GGT AAT GAA CGT CTG**

Reverse primer: **5'-GAA AGC TGG GTG TTA TTT ACG ATG ATC ATC TTC**

*Gsp*DAH7PS^{DAH7PS} *primers* (Round 1)

Forward primer: **5'-GGC AGC GGC GCG ATG GCA CTG CTG GTT AGC**

Reverse primer: **5'- GAA AGC TGG GTG TTA CTC GAG TGC ACG AAC**

Generic primers (round 2)

Forward primer: **5'GGG GAC AAG TTT GTA CAA AAA AGC AGG CTT CGA AAA CCT GTA TTT TCA GGG CAG CGG CGC G**

Reverse primer: **5'GGG GAC CAC TTT GTA CAA GAA AGC TGG GT**

Each PCR mixture contained a total volume of 50 µL consisting of approximately 39.5 ng/µL of PET-28b *Gsp*DAH7PS^{WT} plasmid DNA, 0.5 µM of round one primers, 200 µM of each of the deoxyribonucleotide triphosphates (dNTPs), 10 µL of 5x iProof HF Buffer and 0.02 U/µL iProof DNA polymerase. The remaining volume was made up using autoclaved milliQ water.

Once round one PCR successfullyamplified *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{DAH7PS}, the PCR products were used as the templates for round two of nested PCR using the generic round two primers. Round two PCR used the same concentrations of all the components of round one PCR. Round two PCR for *Gsp*DAH7PS^{DAH7PS} contained approximately 12 ng of the amplified product from round one PCR while round two PCR for *Gsp*DAH7PS^{CM} truncation variant consisted of approximately 19 ng of the amplified product from round one PCR.

Round one PCR of the thermo-cycling program consisted of an initial denaturation at 98 °C for thirty sec followed by 35 cycles of denaturation for ten sec, primer annealing at 45 °C for *Gsp*DAH7PS^{DAH7PS} gene and 50 °C for *Gsp*DAH7PS^{CM} gene for thirty sec with extension at 72 °C (thirty sec for *Gsp*DAH7PS^{DAH7PS} and fifteen sec for *Gsp*DAH7PS^{CM}. Round two PCR followed the same protocol for both variants, the only difference was that the annealing temperature of the PCR reaction for *Gsp*DAH7PS^{CM} truncation variant was 45 °C.

PCR products were treated with Dpn1, 1 µL was used for each reaction mixture and incubated at 37 °C for one hour and purified using E-Gel® CloneWell™ Gel.

5.3.2 Agarose gel electrophoresis

In order to determine if PCR was successful, an agarose gel had to be prepared. The gel separates DNA fragments based on size and this in turn shows if a PCR product with the expected number of base pairs was formed. The gel was composed of 1 % Seakem® LE agarose, 39.2 mL milliQ and 800 µl of 50x tris base-acetic acid-EDTA (TAE) buffer (2 M tris-HCl, 1 M glacial acetic acid, 100 mM EDTA). All components were mixed and heated in a microwave until the agarose had dissolved, then incubated at 60 °C to equilibrate and then then left to cool slightly before adding 4 µL of SYBR® Safe DNA gel stain (Life Technologies). Each sample had 2 µl of 6x loading buffer (60 mM tris-HCl, 60 mM EDTA, 0.02 % (w/v) orange G, 0.05 % (w/v) xylene cyanol FF, 60 % glycerol). The gel was run in a buffer made from 5 mL of 50x TAE buffer mixed with 245 ml milliQ water using a mini-sub cell GT (Bio-Rad) at 85 volts for forty five min. Gels were examined using an Imager® Gel Doc™ XR (Bio-Rad) UV light at 302 nm.

Alternatively, DNA samples were loaded on an E-Gel® 1.2 % precast agarose gel (Life Technologoes) in a buffer-free system and run for twenty seven min. Separation could be monitored in real-time using the E-Gel ® iBase™ Power System and E-Gel® Safe Imager™ Transilluminator at 480 nm. For size comparison, DNA fragments were run alongside a 1 kB Plus DNA ladder (Life Technologies).

5.3.3 BP recombination reaction into pDONR 221

Once a PCR product (contains attB sites added in round two PCR) was obtained, it was ligated into the entry vector pDONR 221 (containing attP sites) through recombination with the use of Gateway® BP Clonase® II enzyme mix. The reaction was carried out by adding cloned DNA segments (approximately 68 ng for *Gsp*DAH7PS^{DAH7PS} and 138 ng for *Gsp*DAH7PS^{CM}), 150 ng of pDONR 221 entry vector and adding enough TE buffer (10 mM tris/HCl, 1 mM EDTA pH 8) to the reaction components to give a final volume of 8µL . The reaction was initiated with the addition of BP clonase, and then incubated at 25 ℃ for one hour before halting the reaction with the addition of proteinase K and incubating the reaction mixture at 37 ℃ for ten min.

5.3.4 Transformation of BP recombination reaction

The BP Recombination Reaction was used to transform competent *E. coli* One Shot® TOP10 Chemically Competent Cells (Life Technologies). A 50 µL aliquot of competent cells was defrosted and incubated on ice for thirty min with 1 µL of the reaction mixture. The cells were then heat shocked for thirty sec at 42 °C and immediately transferred to ice before 250 µL of room temperature super optimal broth (SOC) was added. The cells were incubated at 37 °C for one hour with shaking at 180 rpm before being spread on pre-warmed LB-Kan plates and left to grow overnight at 37 °C.

5.3.5 DNA sequencing

DNA sequencing services were provided by Canterbury Sequencing and Genotyping based at the University of Canterbury. Sequencing was carried out using an Applied Biosystems 3130xl Genetic Analyser and BigDye Terminator v3.1 (Applied Biosystems) used for the sequencing chemistry. Approximately 250 ng of purified double-stranded plasmid and 3.2 µM of primer were supplied for each sample.

5.3.6 LR recombination reaction into destination vectors

Once DNA sequencing results confirmed that the entry vector contained the gene of interest, Gateway® LR Clonase® II enzyme mix was used to catalyse the recombination reaction between pDONR 221 (entry vector containing the gene of interest with attL sites), and the destination vector of choice (pDEST 17-His-tag for *Gsp*DAH7PS^{DAH7PS} and pDEST15-GST-tag for *Gsp*DAH7PS^{CM}) containing the attR sites.

The reaction was carried out by pDONR 221 plasmid (entry vector) containing the gene of interest (approximately 104 ng for *Gsp*DAH7PS^{DAH7PS} and 102 ng for *Gsp*DAH7PS^{CM}), 150 ng of desired destination vector and adding TE buffer (pH 8) to the reaction components to equal 8 µL total volume. The reaction is initiated with the addition of LR clonase, and then incubated a 25 °C for one hour before halting the reaction with the addition of proteinase K and incubating the reaction mixture at 37 °C for ten min.

5.3.7 Transformation of LR recombination reaction

The vectors bearing the genes of *Gsp*DAH7PS^{DAH7PS} and *Gsp*DAH7PS^{CM} were transformed into *E. coli* chaperone 3 cells and *E. coli* BL21 (DE3) cells respectively. 50 µL aliquots of competent cells was defrosted and incubated on ice for thirty min with 1 µL of the reaction mixtures. The cells were then heat-shocked for thirty sec at 42 °C and immediately transferred to ice before 250 µL of room temperature SOC medium was added. The cells were incubated at 37 °C for one hour with shaking (180 rpm) before being spread on pre-warmed LB with the appropriate antibiotics added for *E. coli* BL21(DE3) and *E. coli* chaperone 3 cells. Plates were left to grow overnight at 37 °C.

5.3.8 Plasmid extraction

5 mL cultures were grown for all plasmid extractions. Cultures had the required antibiotics added, inoculated from glycerol stocks and left overnight at 37 °C to grow. Cells were harvested and plasmid extracted using High Pure Plasmid Isolation kit (Roche). The Nanodrop-1000 spectrophotometer was used to determine the concentration of the isolated plasmid at 260 nm.

5.3.9 Transformation

Chemically competent cells (50 µL) were thawed on ice before the addition of 2 µL of either purified PCR product or extracted plasmid. The mixture was left on ice for a further thirty min before the cells were heat-shocked at 42 °C for thirty sec to forty five sec. The cells were then replaced on ice for two min before the addition of 250 µL of SOC medium (2 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgSO4 and 20 mM glucose, filter sterilised) and outgrown at 37 °C for one hour with shaking (200 rpm). An aliquot of transformed cells (75 µL to 200 µL) was spread directly onto a pre-warmed LB-agar plate containing the appropriate antibiotics and left to grow overnight, inverted, at 37 °C.

5.3.10 Glycerol stocks

Once a cell colony carrying a desired gene was obtained, a 5 mL culture was grown from the colony at 37 °C overnight. 750 µL was taken from the culture and mixed with 250 µL of 80 % sterilised glycerol to give a final concentration of 20 % (v/v) in a sterile 1.7 mL microcentrifuge tube, flash frozen with liquid nitrogen and stored at 80 °C.

5.3.11 Media

Cultures were grown in LB (Lennox L) solution made up with milliQ water (20 g/L). LB powder was also used in the making of LB Agar plates used for growing cells. Agar (25 g/L) was mixed with LB powder (20 g/L), and milliQ water and then autoclaved and left to cool. Before using the LB agar, it was heated until all solid was dissolved then incubated at 60 °C for equilibration to occur. Solution was left to cool to avoid antibiotic degradation upon addition and then the solution was poured into Agar plates.

Once the transformation step of *E. coli* cells was carried out, SOC media was added and cells were then incubated at 37 °C for one hour with shaking before being plated.

5.3.12 Purification using nickel affinity column

After cell lysis (for *Gsp*DAH7PS^{WT} and *Gsp*DAH7PS^{DAH7PS}), and the separation of the supernatant from the lysed cells by centrifugation, the supernatant was loaded to a HisTrap HP (GE Healthcare) that had been pre equilibrated with the binding buffer (20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole, pH 7.4), then eluted with the elution buffer (20 mM sodium phosphate, 0.5 M NaCl, 500 mM imidazole, pH 7.4). Eluted fractions of *Gsp*DAH7PS^{DAH7PS} were collected and desalted before the removal of the His-tag using TEV protease. Once TEV protease was added, the mixture was left overnight at 4 °C and then left at room temperature for one hour before being reloaded to a HisTrap HP (GE Healthcare) equilibrated with binding buffe. The protein eluted in the flow through, and was concentrated using a 20 ml 10 000 Da MWCO device (Sartorius Stedim Biotech), and loaded onto a Superdex S200 26/60 SEC column.

5.3.13 Purification using GST affinity chromatography

After cell lysis (for *Gsp*DAH7PS^{CM}), and the separation of the supernatant from the lysed cells by centrifugation, the supernatant was loaded to a GSTrap (GE Healthcare) that had been pre equilibrated with PBS binding buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 pH 7.3) and eluted with the elution buffer (50 mM tris-HCl, 10 mM reduced glutathione, pH 8.0) through a linear gradient. Eluted fractions of *Gsp*DAH7PS^{CM} were collected and buffer exchanged into binding buffer before the removal of the GST-tag using TEV protease. Once TEV protease was added, the mixture was left overnight at 4 °C, then left at room temperature for one hour before being reloaded to a the GSTrap (GE Healthcare) equilibrated with binding buffer. The protein eluted in the flow through, and was concentrated using a 20 ml 5000 kDa MWCO device (Sartorius Stedim Biotech) before being loaded onto a Superdex 75 16/60 SEC column (GE Healthcare).

5.3.14 N-terminal His-tag and GST-tag removal using TEV protease

Fractions containing DAH7PS, as determined by SDS-PAGE, were pooled and washed with loading buffer using a desalting column to remove any imidazole or glutathione. The protein

was then incubated for up to 24 h at 4 °C with recombinant TEV protease (1 OD 280 of TEV protease per 10 OD 280 proteins). The completeness of cleavage was assessed by SDS-PAGE. The rTEV was successfully removed in the next step of purification by reapplication to HisTrap HP or GSTrap column (GE Healthcare).

5.3.15 Protein secondary structure analysis

CD spectra measurements were performed using a JASCO J-815 spectrophotometer. Wavelength scans were carried out in double-distilled water (pH 7.0) between wavelengths 180-240 nm, with 0.5 data pitch, one sec response and 1 mm bandwidth at room temperature. Samples were prepared at an enzyme concentration of 0.15 mg/ml

5.3.16 Metal activation

The metal activation profile of *Gsp*DAH7PS^{DAH7PS} was determined using standard assays with 195 µM PEP and 197 µM E4P in 50 mM BTP buffer Chelex®-treated and filtered, pH 6.8 at 60 °C and 100 µM Cd^{2+}, Mn^{2+}, Mg^{2+}, Zn^{2+}, Ni^{2+}, Co^{2+}, EDTA, DPA or no condition (control). The enzyme (5 µL at 0.13 mg/mL) was incubated and left to equilibrate with the reaction mixture for five min at 60 °C prior to activation with E4P, to a final volume of 1 mL. Triplicate assays were performed.

5.3.17 Temperature dependency study of *Gsp*DAH7PS^{DAH7PS}

The temperature dependency profile of *Gsp*DAH7PS^{DAH7PS} was determined using standard assays. Substrate concentrations were kept constant with 337 µM PEP and 356 µM E4P in 50 mM BTP buffer Chelex®-treated and filtered, pH 6.8 at the desired temperature with 100 µM Cd^{2+}. The reaction mixture was left to equilibrate at the chosen temperature. At each temperature (30 °C, 40 °C, 50 °C, 60 °C, 70 °C and 80 °C), the enzyme, 3 µL at 0.14 mg/mL ($1.4x10^{-2}$ µM in cuvette) was then incubated and left to equilibrate with the reaction mixture for either one or five min prior to activation with E4P, to a final volume of 1 mL. Triplicate assays were performed.

5.3.18 Temperature dependency study of *Gsp*DAH7PS^{CM}

The temperature dependency profile of *Gsp*DAH7PS^{CM} was determined using standard assays. The substrate concentration was kept constant with 80 µM CA in 50 mM BTP buffer Chelex®-treated and filtered, pH 6.8 at the desired temperature. The reaction mixture was left to equilibrate at the chosen temperature. At each temperature (30 °C, 40 °C, 50 °C, 60 °C, 70 °C and 80 °C), the enzyme, 5 µL at 5.29 mg/mL (2.4 µM in cuvette) was then incubated and left to equilibrate with the reaction mixture for either one or five min prior to activation with CA, to a final volume of 1 mL. Triplicate assays were performed.

5.3.19 Michaelis-Menten kinetics for *Gsp*DAH7PS^{DAH7PS}

The assays for determining the kinetic parameters for *Gsp*DAH7PS^{DAH7PS} were carried out in 1 cm path length quartz cuvettes. Standard reaction mixtures contained 100 µM Cd^{2+}, 5 µL protein at 0.11 mg/mL ($1.8x10^{-2}$ µM in cuvette) enzyme in 50 mM BTP buffer (pH 6.8), giving a total volume of 1 mL. To determine the K_m^{E4P}, the PEP concentration was fixed at 332 µM while the E4P concentration was varied (7 µM to 392 µM). For the determination of K_m^{PEP}, E4P concentration was fixed at 324 µM while PEP concentration was varied (12 µM to 354 µM).

5.3.20 Michaelis-Menten kinetics for *Gsp*DAH7PS^{CM}

The assays for determining the kinetic parameters for *Gsp*DAH7PS^{CM} were carried out in 1 mm path length quartz cuvettes. Standard reaction mixtures contained 5 µL enzyme at 0.47 mg/mL (0.84 µM in cuvette) in 50mM BTP buffer (pH 6.8), giving a total volume of 250 µL. To determine the K_m^{CA} different concentrations of CA were used varying from (54 µM to 1632 µM).

5.3.21 CM assay

Steady-state kinetics were measured using a continuous assay monitoring the disappearance of chorismate at 274 nm at 50 °C (extinction coefficient at 274 nm of 2630 M^{-}

^{1} cm^{-1} at 50 °C and pH 6.8) in 50 mM BTP buffer (Chelex®-treated and filtered). Reaction mixtures were made to a final volume of 1 mL and were initiated with the addition of CA at a desired concentration. Initial velocities for the catalytic chorismate mutase activity were measured by initiating the equilibrated reaction mixture with varying concentrations of CA. GraFit (Erithacus Software) was used to fit initial reaction velocities to the Michaelis-Menten equation. An accurate CA substrate concentration was obtained by using a standard assay system. The reaction mixture was made up of enzyme, 50 mM BTP buffer (pH 6.8), enzyme and 10 µL CA, adding up to a total volume of 1 mL. Beer-Lambert's Law was used to determine the concentration of CA ($A = \varepsilon c l$, where $l = 1$ cm and ε of 2630 M^{-1} cm^{-1} at 50 °C) by determining the change in absorbance prior to the addition of enzyme and after the addition of enzyme when the reaction had reached completion (ΔA_1). A control reaction, where CA is absent, was used to determine the change in absorbance caused by the addition of enzyme (ΔA_2). This gave the total change in absorbance ($\Delta A_1 + \Delta A_2$) and was converted to give a more accurate substrate concentration.

5.3.22 Feedback inhibition of *Gsp*DAH7PS^{DAH7PS}

The inhibition profile of *Gsp*DAH7PS^{DAH7PS} was determined using standard assays. Varying concentrations of PA were added to make a 1 mL solution consisting of Chelex®-treated and filtered 50 mM BTP buffer (pH 6.8 at 60 °C), Cd^{2+} (100 µM), PEP (121 µM) and enzyme (7 µM at 0.08 mg/mL). The reaction mixture was incubated at 60 °C for ten min to equilibrate prior to initiating the reaction by addition of E4P (123 µM). All reactions were performed in triplicate.

5.3.23 Analytical SEC

For analytical SEC, the column (Superdex 10/300 GL) was calibrated using known standards to determine the elution volume to mass relationship. Six protein standards; ovalbumin (44 kDa), conalbumin (75 kDa), aldolase (158 kDa), ferritin (440 kDa), thyroglobulin (669 kDa) and blue dextran (2000 kDa), were dissolved in running buffer (same as SEC buffer) and applied to the column. The column was equilibrated with running buffer at a set flow rate of 0.2 mL/min. *Gsp*DAH7PS^{DAH7PS} samples (500 µL volume at 1 mg/mL and 0.5 mg/mL) were

injected onto the equilibrated column and eluted from column using running buffer. After proteins were eluted, the column was equilibrated again for the next run. The peak corresponding to blue dextran was recorded as void volume. The highest point of the protein sample peak was recorded as the elution volume. Elution volumes were recorded and graphed against the log (mass protein standard (Da)). This created a linear graph from which unknown protein masses could be determined.

5.3.24 Analytical ultracentrifugation

Sedimentation velocity and sedimentation equilibrium experiments were performed in a Beckman Coulter Model XL-I analytical ultracentrifuge equipped with UV/Vis scanning optics. Reference buffer solution (10 mM BTP, pH 7.4, 40 mM KCl, 200 µM PEP and 10 µM EDTA) and sample solutions were loaded into 12 mm double-sector cells with quartz windows. The cells were then mounted in an An-60 Ti 4-hole rotor.

Samples at an initial protein concentration of 1.8 mg/mL, 1 mg/mL and 0.5 mg/mL (380 µL) and reference (400 µL) were centrifuged at 50,000 rpm at 20 °C and absorbance data were collected in continuous mode at 278 nm without averaging. Data were fitted to a continuous size distribution model using the program SEDFIT. The partial specific volume (v) of the sample (0.7463 mL g^{-1}), buffer density (1.0004 g mL^{-1}) and buffer viscosity (1.004 cP) were calculated using the program SEDNTERP.

References

1. Sakami, W., and Harrington, H. Amino acid metabolism. *Annual Review of Biochemistry.* **1963,** *32*, 355-398.

2. Young, V. R. Adult amino acid requirements: The case for a major revision in current recommendations. *Journal of Nutrition.* **1994,** *124*, 1517S-1523S.

3. Webby, C. J., Lott, J. S., Baker, H. M., Baker, E. N., and Parker, E. J. Crystallization and preliminary X-ray crystallographic analysis of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Mycobacterium tuberculosis*. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications.* **2005,** *61*, 403-406.

4. Herrmann, K. M., and Weaver, L. M. The shikimate pathway. **1999,** *50*, 473-503.

5. Roberts, F., Roberts, C. W., Johnson, J. J., Kyle, D. E., Krell, T., Coggins, J. R., Coombs, G. H., Milhous, W. K., Tzipori, S., Ferguson, D. J. P., Chakrabarti, D., and McLeod, R. Evidence for the shikimate pathway in apicomplexan parasites. *Nature.* **1998,** *393*, 801-805.

6. Webby, C., Patchett, M., and Parker, E. Characterization of a recombinant type II 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Helicobacter pylori*. *Biochemical Journal.* **2005,** *390*, 223-230.

7. Webby, C. J., Baker, H. M., Lott, J. S., Baker, E. N., and Parker, E. J. The Structure of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from Mycobacterium tuberculosis reveals a common catalytic scaffold and ancestry for type I and type II enzymes. *Journal of Molecular Biology.* **2005,** *354*, 927-939.

8. Shumilin, I. A., Zhao, C., Bauerle, R., and Kretsinger, R. H. Allosteric Inhibition of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate Synthase Alters the Coordination of Both Substrates. *Journal of Molecular Biology.* **2002,** *320*, 1147-1156.

9. Schneider, T. R., Hartmann, M., and Braus, G. H. Crystallization and preliminary X-ray analysis of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (tyrosine inhibitable) from *Saccharomyces cerevisiae*. *Acta Crystallographica Section D: Biological Crystallography.* **1999,** *55*, 1586-1588.

10. Raetz, C. R. Biochemistry of endotoxins. *Annual Review of Biochemistry.* **1990,** *59*, 129-170.

11. Blackmore, N. J., Reichau, S., Jiao, W., Hutton, R. D., Baker, E. N., Jameson, G. B., and Parker, E. J. Three sites and you are out: Ternary synergistic allostery controls aromatic amino acid biosynthesis in mycobacterium tuberculosis. *Journal of Molecular Biology.* **2013,** *425*, 1582-1592.

12. Wu, J., and Woodard, R. W. New insights into the evolutionary links relating to the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase subfamilies. *Journal of Biological Chemistry.* **2006,** *281*, 4042-4048.

13. Wu, J., Sheflyan, G., and Woodard, R. *Bacillus subtilis* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase revisited: resolution of two long-standing enigmas. *Biochemical Journal.* **2005,** *390*, 583-590.

14. Schofield, L. R., Anderson, B. F., Patchett, M. L., Norris, G. E., Jameson, G. B., and Parker, E. J. Substrate ambiguity and crystal structure of *Pyrococcus furiosus* 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase: an ancestral 3-deoxyald-2-ulosonate-phosphate synthase? *Biochemistry.* **2005,** *44*, 11950-11962.

15. Zhou, L., Wu, J., Janakiraman, V., Shumilin, I. A., Bauerle, R., Kretsinger, R. H., and Woodard, R. W. Structure and characterization of the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Aeropyrum pernix*. *Bioorganic Chemistry.* **2012,** *40*, 79-86.

16. Cross, P. J., Dobson, R. C., Patchett, M. L., and Parker, E. J. Tyrosine latching of a regulatory gate affords allosteric control of aromatic amino acid biosynthesis. *Journal of Biological Chemistry.* **2011,** *286*, 10216-10224.

17. Jiao, W., Hutton, R. D., Cross, P. J., Jameson, G. B., and Parker, E. J. Dynamic cross-talk among remote binding sites: The molecular basis for unusual synergistic allostery. *Journal of Molecular Biology.* **2012,** *415*, 716-726.

18. Parker, E. J., Bulloch, E. M., Jameson, G. B., and Abell, C. Substrate deactivation of phenylalanine-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase by erythrose 4-phosphate. *Biochemistry.* **2001,** *40*, 14821-14828.

19. Jensen, R. A., Xie, G., Calhoun, D. H., and Bonner, C. A. The correct phylogenetic relationship of KdsA (3-deoxy-D-*manno*-octulosonate 8-phosphate synthase) with one of two independently evolved classes of AroA (3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase). *Journal of Molecular Evolution.* **2002,** *54*, 416.

20. Subramaniam, P. S., Xie, G., Xia, T., and Jensen, R. A. Substrate ambiguity of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* in the context of its membership in a protein family containing a subset of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthases. *Journal of Bacteriology.* **1998,** *180*, 119-127.

21. Walker, G. E., Dunbar, B., Hunter, I. S., Nimmo, H. G., and Coggins, J. R. Evidence for a novel class of microbial 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase in *Streptomyces coelicolor* A3(2), *Streptomyces rimosus* and *Neurospora crassa*. *Microbiology.* **1996,** *142*, 1973-1982.

22. Oliynyk, Z., Briseno-Roa, L., Janowitz, T., Sondergeld, P., and Fersht, A. Designing a metal-binding site in the scaffold of *Escherichia coli* KDO8PS. *Protein Engineering Design and Selection.* **2004,** *17*, 383-390.

23. Ahn, M., Pietersma, A. L., Schofield, L. R., and Parker, E. J. Mechanistic divergence of two closely related aldol-like enzyme-catalysed reactions. *Organic and Biomolecular Chemistry.* **2005,** *3*, 4046-4049.

24. Ray, P. H., and Benedict, C. D. Purification and characterization of a specific 3-deoxy-D-*manno*-octulosonate 8-phosphate phosphatase from *Escherichia coli B*. *Journal of Bacteriology.* **1980,** *142*, 60-68.

25. Staub, M., and Dénes, G. A kinetic study of the mechanism of action of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase in *Escherichia coli* K -12. *Biochimica et Biophysica Acta (BBA)-Enzymology.* **1967,** *132*, 528-530.

26. DeLeo, A. B., Dayan, J., and Sprinson, D. Purification and kinetics of tyrosine-sensitive 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate synthetase from *Salmonella*. *Journal of Biological Chemistry.* **1973,** *248*, 2344-2353.

27. Kohen, A., Berkovich, R., Belakhov, V., and Baasov, T. Stereochemistry of the KDO8P synthase. An efficient synthesis of the 3-fluoro analogues of KDO8P. *Bioorganic & Medicinal Chemistry Letters.* **1993,** *3*, 1577-1582.

28. Onderka, D., and Floss, H. Steric course of the chorismate synthetase reaction and the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthetase reaction. *Journal of the American Chemical Society.* **1969,** *91*, 5894-5896.

29. Reichau, S., Jiao, W., Walker, S. R., Hutton, R. D., Baker, E. N., and Parker, E. J. Potent Inhibitors of a Shikimate Pathway Enzyme from *Mycobacterium tuberculosis* combining mechanism and modeling based design. *Journal of Biological Chemistry.* **2011,** *286*, 16197-16207.

30. Lambert, J. M., Boocock, M. R., and Coggins, J. R. The 3-dehydroquinate synthase activity of the pentafunctional arom enzyme complex of *Neurospora crassa* is Zn2+-dependent. *Biochemical Journal.* **1985,** *226*, 817-829.

31. Paravicini, G., Schmidheini, T., and Braus, G. Purification and properties of the 3‐deoxy‐D‐*arabino*‐heptulosonate‐7‐phosphate synthase (phenylalanine‐inhibitable) of *Saccharomyces cerevisiae*. *European Journal of Biochemistry.* **1989,** *186*, 361-366.

32. Ray, J., and Bauerle, R. Purification and properties of tryptophan-sensitive 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*. *Journal of Bacteriology.* **1991,** *173*, 1894-1901.

33. Krosky, D. J., Alm, R., Berg, M., Carmel, G., Tummino, P. J., Xu, B., and Yang, W. *Helicobacter pylori* 3-deoxy-D- *manno*-octulosonate-8-phosphate (KDO-8-P) synthase is a zinc-metalloenzyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology.* **2002,** *1594*, 297-306.

34. Tao, P., Schlegel, H. B., and Gatti, D. L. Common basis for the mechanism of metallo and non-metallo KDO8P synthases. *Journal of Inorganic Biochemistry.* **2010,** *104*, 1267-1275.

35. Duewel, H. S., and Woodard, R. W. A metal bridge between two enzyme families: 3-Deoxy-D-*manno*-octulosonate-8-phosphate synthase from *Aquifex aeolicus* requires a divalent metal for activity. *Journal of Biological Chemistry.* **2000,** *275*, 22824-22831.

36. Stephens, C. M., and Bauerle, R. Analysis of the metal requirement of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*. *Journal of Biological Chemistry.* **1991,** *266*, 20810-20817.

37. Shumilin, I. A., Kretsinger, R. H., and Bauerle, R. H. Crystal structure of phenylalanine-regulated 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*. *Structure.* **1999,** *7*, 865-875.

38. Wu, J., Howe, D. L., and Woodard, R. W. *Thermotoga maritima* 3-deoxy-D-*arabino*-heptulosonate 7-Phosphate (DAHP) Synthase. *Journal of Biological Chemistry.* **2003,** *278*, 27525-27531.

39. Cross, P. J., Pietersma, A. L., Allison, T. M., Wilson‐Coutts, S. M., Cochrane, F. C., and Parker, E. J. *Neisseria meningitidis* expresses a single 3‐deoxy‐D‐*arabino*‐heptulosonate 7‐phosphate synthase that is inhibited primarily by phenylalanine. *Protein Science.* **2013,** *22*, 1087-1099.

40. Schofield, L. R., Patchett, M. L., and Parker, E. J. Expression, purification, and characterization of 3-deoxy-D-*arabino*- heptulosonate 7-phosphate synthase from *Pyrococcus furiosus*. *Protein Expression and Purification.* **2004,** *34*, 17-27.

41. Shumilin, I. A., Bauerle, R., Wu, J., Woodard, R. W., and Kretsinger, R. H. Crystal structure of the reaction complex of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate Synthase from *Thermotoga maritima* refines the catalytic mechanism and indicates a new mechanism of allosteric regulation. *Journal of Molecular Biology.* **2004,** *341*, 455-466.

42. Wierenga, R. The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS letters.* **2001,** *492*, 193-198.

43. Hu, C., Jiang, P., Xu, J., Wu, Y., and Huang, W. Mutation analysis of the feedback inhibition site of phenylalanine‐sensitive 3‐deoxy‐D‐*arabino*‐heptulosonate 7‐phosphate synthase of *Escherichia coli*. *Journal of Basic Microbiology.* **2003,** *43*, 399-406.

44. Grant, G. A. The ACT domain: a small molecule binding domain and its role as a common regulatory element. *Journal of Biological Chemistry.* **2006,** *281*, 33825-33829.

45. Aravind, L., and Koonin, E. V. Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. *Journal of Molecular Biology.* **1999,** *287*, 1023-1040.

46. Webby, C. J., Jiao, W., Hutton, R. D., Blackmore, N. J., Baker, H. M., Baker, E. N., Jameson, G. B., and Parker, E. J. Synergistic allostery, a sophisticated regulatory network for the control of aromatic amino acid biosynthesis in *Mycobacterium tuberculosis*. *Journal of Biological Chemistry.* **2010,** *285*, 30567-30576.

47. König, V., Pfeil, A., Braus, G. H., and Schneider, T. R. Substrate and Metal Complexes of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate Synthase from *Saccharomyces cerevisiae* provide new insights into the catalytic mechanism. *Journal of Molecular Biology.* **2004,** *337*, 675-690.

48. Tribe, D., Camakaris, H., and Pittard, J. Constitutive and repressivle enzymes of the common pathway of aromatic biosynthesis in *Escherichia coli* K-12: regulation of enzyme synthesis at different growth rates. *Journal of Bacteriology.* **1976,** *127*, 1085-1097.

49. Hartmann, M., Schneider, T. R., Pfeil, A., Heinrich, G., Lipscomb, W. N., and Braus, G. H. Evolution of feedback-inhibited/barrel isoenzymes by gene duplication and a single mutation. *Proceedings of the National Academy of Sciences.* **2003,** *100*, 862-867.

50. Cross, P. J., Allison, T. M., Dobson, R. C., Jameson, G. B., and Parker, E. J. Engineering allosteric control to an unregulated enzyme by transfer of a regulatory domain. *Proceedings of the National Academy of Sciences.* **2013,** *110*, 2111-2116.

51. Tzin, V., Malitsky, S., Zvi, M. M. B., Bedair, M., Sumner, L., Aharoni, A., and Galili, G. Expression of a bacterial feedback‐insensitive 3‐deoxy‐D‐*arabino*‐heptulosonate 7‐phosphate synthase of the shikimate pathway in *Arabidopsis elucidates* potential metabolic bottlenecks between primary and secondary metabolism. *New Phytologist.* **2012,** *194*, 430-439.

52. Cotton, R. G., and Gibson, F. The biosynthesis of phenylalanine and tyrosine; enzymes converting chorismic acid into prephenic acid and their relationships to prephenate dehydratase and prephenate dehydrogenase. *Biochimica et Biophysica Acta (BBA)-General Subjects.* **1965,** *100*, 76-88.

53. Lee, A. Y., Karplus, P. A., Ganem, B., and Clardy, J. Atomic structure of the buried catalytic pocket of *Escherichia coli* chorismate mutase. *Journal of the American Chemical Society.* **1995,** *117*, 3627-3628.

54. Xue, Y., and Lipscomb, W. N. Location of the active site of allosteric chorismate mutase from *Saccharomyces cerevisiae*, and comments on the catalytic and regulatory mechanisms. *Proceedings of the National Academy of Sciences.* **1995,** *92*, 10595-10598.

55. Ökvist, M., Dey, R., Sasso, S., Grahn, E., Kast, P., and Krengel, U. 1.6 Å Crystal structure of the secreted chorismate mutase from *Mycobacterium tuberculosis*: Novel fold topology revealed. *Journal of Molecular Biology.* **2006,** *357*, 1483-1499.

56. Gu, W., Williams, D. S., Aldrich, H. C., Xie, G., Gabriel, D. W., and Jensen, R. A. The AroQ and PheA domains of the bifunctional P-protein from *Xanthomonas campestris* in a context of genomic comparison. *Microbial & Comparative Genomics.* **1997,** *2*, 141-158.

57. Romero, R., Roberts, M., and Phillipson, J. Chorismate mutase in microorganisms and plants. *Phytochemistry.* **1995,** *40*, 1015-1025.

58. Chook, Y. M., Ke, H., and Lipscomb, W. N. Crystal structures of the monofunctional chorismate mutase from *Bacillus subtilis* and its complex with a transition state analog. *Proceedings of the National Academy of Sciences.* **1993,** *90*, 8600-8603.

59. Helmstaedt, K., Heinrich, G., Merkl, R., and Braus, G. H. Chorismate mutase of *Thermus thermophilus* is a monofunctional AroH class enzyme inhibited by tyrosine. *Archives of Microbiology.* **2004,** *181*, 195-203.

60. Gray, J. V., Golinelli-Pimpaneau, B., and Knowles, J. R. Monofunctional chorismate mutase from *Bacillus subtilis*: purification of the protein, molecular cloning of the gene, and overexpression of the gene product in *Escherichia coli*. *Biochemistry.* **1990,** *29*, 376-383.

61. Light, S. H., Halavaty, A. S., Minasov, G., Shuvalova, L., and Anderson, W. F. Structural analysis of a 3‐deoxy‐D‐*arabino*‐heptulosonate 7‐phosphate synthase with an N‐terminal chorismate mutase‐like regulatory domain. *Protein Science.* **2012,** *21*, 887-895.

62. Kim, S. K., Reddy, S. K., Nelson, B. C., Robinson, H., Reddy, P. T., and Ladner, J. E. A comparative biochemical and structural analysis of the intracellular chorismate mutase (Rv0948c) from *Mycobacterium tuberculosis* H(37)R(v) and the secreted chorismate mutase (y2828) from *Yersinia pestis*. *FEBS Journal.* **2008,** *275*, 4824-4835.

63. Sasso, S., Ökvist, M., Roderer, K., Gamper, M., Codoni, G., Krengel, U., and Kast, P. Structure and function of a complex between chorismate mutase and DAHP synthase: efficiency boost for the junior partner. *The EMBO Journal.* **2009,** *28*, 2128-2142.

64. Topanurak, S., Sinchaikul, S., Sookkheo, B., Phutrakul, S., and Chen, S. T. Functional proteomics and correlated signaling pathway of the thermophilic bacterium *Bacillus stearothermophilus* TLS33 under cold‐shock stress. *Proteomics.* **2005,** *5*, 4456-4471.

65. Nazmi, A. R. Personal communication.

66. Lea, W. A., and Simeonov, A. Differential scanning fluorometry signatures as indicators of enzyme inhibitor mode of action: case study of glutathione S-transferase. *PloS one.* **2012,** *7*, e36219.

67. Tran, D., Pietersma, A. L., Schofield, L. R., Rost, M., Jameson, G. B., and Parker, E. J. Investigating the role of the hydroxyl groups of substrate erythrose 4-phosphate in the reaction catalysed by the first enzyme of the shikimate pathway. *Bioorganic & Medicinal Chemistry Letters.* **2011,** *21*, 6838-6841.

68. Cross, P. J., and Parker, E. J. Allosteric inhibitor specificity of Thermotoga maritima 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *FEBS letters.* **2013,** *587*, 3063-3068.

69. Zhang, S., Kongsaeree, P., Clardy, J., Wilson, D. B., and Ganem, B. Site-directed mutagenesis of monofunctional chorismate mutase engineered from the *Escherichia coli* P-protein. *Bioorganic & Medicinal Chemistry.* **1996,** *4*, 1015-1020.

70. Nazmi, A. R., Schofield, L. R., Dobson, R. C., Jameson, G. B., and Parker, E. J. Destabilization of the Homotetrameric Assembly of 3-Deoxy-D-*arabino*-Heptulosonate-7-Phosphate Synthase from the Hyperthermophile *Pyrococcus furiosus* Enhances Enzymatic Activity. *Journal of Molecular Biology.* **2014,** *426*, 656-673.

71. Schoner, R., and Herrmann, K. M. 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. Purification, properties, and kinetics of the tyrosine-sensitive isoenzyme from *Escherichia coli*. *Journal of Biological Chemistry.* **1976,** *251*, 5440-5447.

72. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical Journal.* **2000,** *78*, 1606-1619.

73. Lebowitz, J., Lewis, M. S., and Schuck, P. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Science.* **2009,** *11*, 2067-2079.